On quantum computing for artificial superintelligence

Anna Grabowska[1,2] · Artur Gunia[2]

Abstract

Artificial intelligence algorithms, fueled by continuous technological development and increased computing power, have proven effective across a variety of tasks. Concurrently, quantum computers have shown promise in solving problems beyond the reach of classical computers. These advancements have contributed to a misconception that quantum computers enable hypercomputation, sparking speculation about quantum supremacy leading to an intelligence explosion and the creation of superintelligent agents. We challenge this notion, arguing that current evidence does not support the idea that quantum technologies enable hypercomputation. Fundamental limitations on information storage within finite spaces and the accessibility of information from quantum states constrain quantum computers from surpassing the Turing computing barrier. While quantum technologies may offer exponential speed-ups in specific computing cases, there is insufficient evidence to suggest that focusing solely on quantum-related problems will lead to technological singularity and the emergence of superintelligence. Subsequently, there is no premise suggesting that general intelligence depends on quantum effects or that accelerating existing algorithms through quantum means will replicate true intelligence. We propose that if superintelligence is to be achieved, it will not be solely through quantum technologies. Instead, the attainment of superintelligence remains a conceptual challenge that humanity has yet to overcome, with quantum technologies showing no clear path toward its resolution.

✉ Anna Grabowska
anna.belowska@doctoral.uj.edu.pl

Artur Gunia
artur.gunia@uj.edu.pl

1 Doctoral School of the Social Sciences, Jagiellonian University, Rynek Glowny 34, Krakow 31-010, Poland

2 Institute of Philosophy, Jagiellonian University, ul. Grodzka 52, Krakow 33-332, Poland

1 Introduction

It is widely maintained that humanity today is united into one entity whose brain represents the collective knowledge of humanity, with a collective intelligence far greater than that of the individual (see Weschsler, 1971; Muthukrishna & Henrich, 2016; Mulgan, 2018). According to Ray Kurzweil, the continuous development and growth of knowledge follows an exponential curve, culminating in a point known as technological singularity. Crossing this threshold will substantially change human reality (see Kurzweil, 2005). The concept of singularity, understood as a technological breakthrough that brings unforeseeable changes, is not new and was already present in the middle of the twentieth century in the views of John von Neuman, the father of modern computer science (see Ulam et al., 2013). Technological singularity is frequently linked or even identified with the concept of an intelligence explosion, a hypothetical scenario in which an artificial intelligence system starts to improve its own intelligence in a recursive manner (e.g., Good, 1966). This self-improving cycle yields an artificial super-intelligent agent that surpasses humans in intelligence *quantitatively* as well as *qualitatively* and thus possesses unimaginable cognitive and knowledge boundaries. The emergence of a general form of intelligence that exceeds human understanding and control will be the moment of reaching the technological singularity. Thus, since the 1960s, the emergence of artificial general intelligence (AGI) has been closely linked to knowledge growth and technological development.

However, Kurzweil solidified this idea by tying the exponential curve of technological development to Moore's law and computational capacity (Kurzweil, 2005). Moore's law states that transistors on a microprocessor, and thus computational power, double approximately every two years (Moore, 1965; Kurzweil, 2005). Since modern processors already consist of billions of transistors, the perspective of an exponential increase in the number of transistors on a microprocessor raised reasonable expectations for an increase in computational power large enough to cross technological singularity (Chalmers, 2016; Bostrom, 2014; Müller & Bostrom, 2016). However, the increase in computational power in processors has slowed down recently, as the technological limits of packaging transistors on such small surfaces as microprocessors are being reached (Mack, 2011). Consequently, quantum technologies that are unburdened by the space limitations of transistors have gained widespread interest as potential successors to transistors capable of sustaining exponential growth in computational capacity (Möller & Vuik, 2017).

Indeed, to date, quantum computers, compared to classic machines, have already shown outstanding improvements in a variety of computational problems that demand an extensive amount of computing resources, such as integer factorization, combinatorial optimization, differential equation solvers, many-body systems simulations; they also have shown a significant acceleration in various artificial intelligence (AI) algorithms (for a review, see Bharti et al., 2022). In 1998, Nick Bostrom linked superintelligence development with an increase in computational power driven by the development of quantum technology (Bostrom,

1998). Later, the opinion that quantum computers will solve problems that no classical computer could solve (called *quantum supremacy*; Preskill, 2012) and thus will contribute to the intelligence explosion appeared in the works of many AI researchers. In an essay entitled *Artificial General Intelligence and the Future of Humanity*, Ben Goertzel expressed the conviction that singularity is going to involve a variety of different technologies, including quantum computing (Goertzel, 2013). Similar statements can be found in the works of Callaghan et al. (2017) and Zheng and Akhmad (2017). A prediction of the forthcoming singularity, accelerated by quantum computers, can also be found in recent works by Yampolskiy (2018), Miller (2019), and Baaquie and Kwek (2023).

While quantum computers have brought solutions to many previously unsolvable problems (Aaronson, 2010), there is ongoing debate whether they can truly surpass the computational limits of classical machines outlined by mathematics and solve all the problems considered unsolvable, i.e. perform *hypercomputations*. The idea that quantum computers can perform hypercomputations was introduced for the first time in 1946 by Arthur Komar (Salisbury, 2022). Since then, there have been numerous works describing theoretical systems capable of hypercomputation (for a review, see Copeland, 2002; see also Copeland, 2004; Copeland & Shagrir, 2011; Shagrir & Pitowsky, 2003). These works, however, met with strong objections (for a range of viewpoints, see Davis, 2004; Hodges, 2005; Hagar and Korolev, 2006, 2007; Cotogno, 2009). Most recently, Scott Aaronson aimed to demonstrate a contradiction between modern physics and claims suggesting that quantum computers may perform hypercomputations (Aaronson et al., 2022). Although researchers when discussing the possibility of superintelligence usually do not refer to the hypercomputation problem (see Aaronson, 2013b), the discussion on hypercomputation is central to the debate about the contribution of quantum computers to the technological singularity and the intelligence explosion.

The aim of this study is to show, based on the discourse surrounding hypercomputability, that quantum technologies do not enable the development of artificial superintelligence. Our argument is as follows: from the beginning of computing machines, the creation of artificial superintelligence has been tied to technological development and to increase in computational power (Good, 1966; Kurzweil, 2005). The computational power of classic computing machines proved to be limited (Llyod, 2000; Markov, 2014). Several attempts have been made to prove that quantum computers can surpass the limitations of the computing power of classical computing machines (Copeland, 2002), and therefore they might facilitate the creation of artificial superintelligence (Bostrom, 2014; Müller & Bostrom, 2016). However, various mathematical theories show that there are hard physical boundaries that limit the growth of computational power, regardless of the computing machine (Davis, 2004). Thus, assuming that the created mathematical theories are correct approximations of the observable world, and that the creation of artificial superintelligence is linked with an increase in computational power, quantum computers cannot contribute to the creation of artificial superintelligence. To expand the discussion of the role of quantum technologies in achieving superintelligence beyond the most popular path involving technological singularity, we explore and discuss two alternative paths in which quantum computers could contribute to the emergence

of superintelligence. These paths involve the significant acceleration of existing AI algorithms (Moret-Bonillo, 2015; Liu et al., 2021; but see also Wiedermann, 2012) and the emergence of the quantum mind, with its medium being a quantum computer (Penrose, 1994; Mikki, 2023). We show that, as of today, there is no evidence that AI algorithms are close to achieving general intelligence, nor is there evidence that the emergence of either mind or intelligence is related to quantum effects.

2 Perspectives of superintelligence

Intelligence is usually defined as the ability to understand, and use acquired knowledge in new situations (Pintner, 1923). It also frequently refers to the effectiveness of activities that require cognition (Herrnstein & Murray, 2010). Therefore, intelligence is one of the most interdisciplinary aspects of the human mental state (Sternberg, 1985). Philosophy tries to answer what it means to be intelligent; psychology and cognitive science focus on the purpose of intelligence and how to use it effectively in problem-solving; neuroscience studies the structure of the biological, chemical, and physical features responsible for rationality and intelligent behavior. These disciplines broaden our knowledge of intelligence, although many of the questions related to it have not yet been satisfactorily answered. Despite this lack of understanding of the nature of intelligence, humanity desires both to improve its own intellect and to bestow it on artificial creatures (More, 2013), which in the field of computer science manifests itself in AI.

Turing (1950) reflected on the mathematical prospects of AI, discussing how to build intelligent machines and how to test their intelligence. The actual consideration of AI began with a research project at Dartmouth Workshop in 1956 by John McCarthy, Marvin L. Minsky, Nathaniel Rochester and Claude E. Shannon, who described the AI problem as follows: 'If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have (McCarthy et al., 2006). One of the recent views has defined AI as systems that think and behave like humans, i.e. think and behave rationally (Russell et al., 2003).

Today, AI manifests itself as numerous practical solutions, e.g. expert systems, neural networks, machine learning, deep learning, computer vision, natural language processing, and many more (Flasiński, 2016). These solutions fall into the category of narrow or weak AI, meaning that they only mimic a particular mental process. Narrow AI systems can enhance human intelligence by supporting the calculations, analyses, or pattern recognition performed in the brain. The opposite of narrow AI is strong AI, deep AI, or AGI. John Searle explained it as follows: in 'strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states' (Searle, 1980).

Although AGI is still absent in the current world, it is a phenomenon that is interesting from many perspectives: theory of mind, moral abilities, or self-awareness.

From a first point of view, an AI system should interact with society and be able to understand emotions, thoughts, and expectations. Factually, modern AI algorithms for affective computing have been shown to recognize (often better than humans), interpret, process, and simulate human emotions (Picard, 2000). Moral modules are beginning to be present in today's applications and embodied agents like robots are present as well (Indurkhya, 2019; Gunia et al., 2024). The perspective of self-awareness requires AI agents to be self-reflective and sentient beings, which currently remains in the realm of science fiction. However, such a development might emerge from the web of many narrow AI systems.

AGI is a fairly general and vague term. In the context of intelligent machines, the more specific term, which is gaining in popularity (especially in the transhumanist movement), is superintelligence. Nick Bostrom defined superintelligence as 'any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest' (Bostrom, 2014), or 'an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills' (Bostrom, 1998). Within the transhumanist movement, there is a conviction that AI software will evolve into superintelligence and that superintelligence will teach humanity not only science and technology but also spiritual and moral principles (Lee, 2019). The problem of superintelligence is usually associated with one of the following conceptions: designing human-independent intelligence; the development of a superintelligent environment; and the technological improvement of human mental abilities. In particular, according to Bostrom (2014) and many transhumanists, the emergence of superintelligence can be considered from several perspectives: cognitive enhancement, collective superintelligence, mind upload, and artificial superintelligence.

Cognitive enhancement is the act of increasing or extending the core abilities of the brain or mind by improving or adding to its internal or external information processing systems (Sandberg & Bostrom, 2006). The term encapsulates the improvement of intelligence and attention, the reinforcement of creativity and memory, and the extension of perception range. Cognitive enhancement also covers the issues of improving control, emotional coping, or motivation (Kipke, 2013). Examples of cognitive enhancements include pharmacological neuroenhancement, i.e., the use of psychoactive substances to enhance vigilance, concentration, memory or mood by healthy individuals; genetic enhancements, i.e., the use of genetic engineering to change physical appearance, metabolism or improve physical and mental abilities (Kutt et al., 2015); and the use of cognitive technology that actively affects human cognitive processes (Gunia & Indurkhya, 2017). All processes intended to result in enhancement must be voluntary and go beyond the natural and nonpathological limitations of human cognition (Hauskeller, 2014). However, these perspectives seem to be time-consuming and ineffective, considering the main assumptions of superintelligence.

Collective superintelligence is a form of group intelligence that emerges from the gradual improvement of networks and organizations that link individual human minds together with various artifacts. Instead of a single individual, an entire system of properly organized and networked individuals might attain a form of superintelligence (Bostrom, 2014). This idea includes a broad vision

of the networked devices connected to the brain. Ray Kurzweil noted his belief that nanobots, nanosized robots, will inhabit human bodies, keep them healthy, and constantly transmit human brains into the collective cloud (Kurzweil, 2005). Understood in this way, superintelligence seems easily achievable, but such a solution would require computation on an unprecedented and probably unattainable scale.

Another way superintelligence may emerge is mind uploading or whole-brain emulation. Mind uploading is the process of transferring the mind from a biological brain to another substrate (Koene, 2013). The process involves scanning the physical structure of the brain to create an emulation of a mental state, including long-term memory and human self-identity. The data obtained are then transferred or copied to a digital format (Goertzel & Ikle, 2012). Such an approach requires combining knowledge from various disciplines: neuroinformatics, neurobiology, nanotechnology, virtual reality, and philosophy of the mind, to copy every function of each neuron and the more general mind functions that emerge from neuron interactions. The key factor within this idea is the medium – hardware and software – into which the mind will be transferred. It is unknown whether the capabilities of classic computers based on transistor processors are sufficient to emulate the mind, or whether mind uploading requires a different solution, including, for example, quantum computers.

Superintelligence usually refers to artificial superintelligence. Bostrom (2014) noted that artificial superintelligence starts with a computer that is marginally smarter than a human and ends with an AI system a billion times smarter. Good Old Fashioned Artificial Intelligence Systems, in his opinion, did not focus on learning, uncertainty, or concept formation, perhaps due to technical limitations. In contrast, the ability to learn would be an essential feature of the system that aims to attain artificial superintelligence. It should be emphasized that AI does not have to resemble the human mind and that artificial superintelligence can thus develop on different principles than human intelligence.

Bostrom (2014) highlights different forms of superintelligence: (1) speed superintelligence: a system that can perform all the cognitive tasks that the human intellect can, but much faster. It would be the simplest form of brain emulation that exceeds human intelligence quantitatively but not qualitatively; (2) collective superintelligence: a system composed of a large number of smaller intellects in such a way that the overall performance of the system across many general domains is far superior to all current cognitive systems. Such a form of superintelligence would bind various intelligent structures, both human and non-human, and could be constantly improved with new components. Nevertheless, as mentioned before, collective superintelligence seems difficult to achieve; (3) quality superintelligence: a system that is both quantitatively and qualitatively superior. It is a form of superintelligence that goes beyond human capacity, combining elements of animal intelligence with other cognitive abilities that are currently incomprehensible. Like in the previous case, the creation of such a system would require enormous computing power, which is not available right now. Although Bostrom (2014) did not focus on the technical aspects, he expressed his conviction that superintelligence will be realized digitally, through both hardware and software advantages of forms capable of running on computers

Table 1 Hardware and software advantages of non-biological intelligence according to Bostrom, 2014

Non-biological intelligence advantage	Manifestation	Example
Hardware advantages	Speed of internal communication and information flow.	Data transmission is close to the speed of light, while the peak speed of neurons is much slower.
	The number of computational elements.	The brain has a limited number of elements, while computers can be constantly upgraded.
	Storage capacity.	Human working memory can hold no more than five chunks of information at any given time, while computers can store much more information at once.
Software advantages	Editability.	Digital structures can be manipulated, experimented and tweaked, while such experiments on biological brains are difficult to perform, with minimal effects and various ethical questions.
	Duplicability.	A computer program can be backed up any number of times.
	Goal coordination.	Computers, unlike humans, go straight to the goal without deviating from it, being essentially goal-oriented.
	Memory sharing.	Biological brains need a long period of training and education under the guidance of mentors, while digital minds can acquire new memories and skills by exchanging data files.

(see. Table 1.). In conclusion, artificial superintelligence is supposed to be fast and flexible and it is assumed to be able to conclude a new quality.

Superintelligence raises many doubts. It is unclear whether its actions will be sufficiently explainable to humans or whether they will become a factor of existential risk (Bostrom, 2013). It is an open question of which medium will enable the implementation of superintelligence, assuming its most probable source is the AGI. Both classical computers and the computing power of the brain have limitations, and the combination of these two mediums is difficult to implement. Even if the functionality of the biological and digital mind could be similar, the physical structure and main phenomena are different. With recent advances in quantum physics, and thus in quantum computing, the question arises as to whether quantum computers can be used to implement the properties attributed to superintelligence. In the following sections, we investigate this issue.

3 Quantum computers

Despite the name, quantum computers have little in common with computers used every day. Their foundations were established in the 1980s by Paul Benioff, Richard Feynmann, and David Deutsch (Benioff, 1982; Deutsch, 1985; Feynman, 1982), but recent latest technological advancements have brought an intensive acceleration in the development of quantum computing technologies (for a review, see Ladd et al., 2010).

3.1 Principles of quantum computing: qubit, superposition, and entanglement

The fundamental unit of information in quantum information processing is the qubit, a quantum analog of the bit in classical computers. A bit is a logical unit that represents the state of a system; this logical unit can take one of two values, usually referred to as 0 and 1. These values could be physically realized in various ways, e.g., through a voltage change. Just as principles of classical physics underlie the state of a bit, quantum physics is responsible for the state of a qubit. In contrast to classical physics, in quantum physics the state of a system is undetermined until it is measured; thus, the state of a system before measuring its state is expressed in terms of the probability of being in each possible state. As a result, the qubit has no value of 0 or 1, but is a linear combination, that is, a *superposition*, of the basis states $|0\rangle$ and $|1\rangle$, where the states $|0\rangle$ and $|1\rangle$ are quantum equivalents of states 0 and 1 from classical physics. This probability of being in each possible state can be easily expressed with a linear function $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where ψ is a wave of probability that a particle will be in a given state after measurement, and coefficients α and β are complex numbers that represent probabilities of this particle being in $|0\rangle$ and $|1\rangle$, respectively. Such an intuitive representation of a complex phenomenon was proposed by Erwin Schrödinger in 1926 and later refined by Max Born (Born & Wiener, 1926). Since α and β denote probabilities, they are subject to the law of total

probability and must fulfill the following requirement: $|\alpha|^2 + |\beta|^2 = 1$ (see *quantum law of total probability*; Yang, 2022). Physically, the qubit can be realized by any system that is quantum and has two different basis states, e.g. two energy levels in an atom.

The design of a system's state in the form of superposition underlies one of the most often mentioned advantages of quantum over classical computing (Swami & Gill, 2011; Avaliani, 2004). In 1946 Felix Bloch proved that when α and β fulfill the requirements of the quantum law of total probability, the qubit can be represented as a point on a three-dimensional complex sphere called the Bloch sphere (Bloch, 1946). Since there is an infinite number of points on the Bloch sphere (precisely, on any surface), theoretically a single qubit can store an infinite amount of information (see Nielsen & Chuang, 2010; Benenti et al., 2004). Thus, the phenomenon of superposition became the foundation for several propositions of quantum hypercomputation (e.g., Deutsch, 1985; Li et al., 2001). Another important characteristic of quantum states is *quantum entanglement.* Qubits can be assembled into complex particle systems that exhibit an important property called quantum entanglement (Nielsen & Chuang, 2010). In an entangled system, the state of the system is always expressed as a superposition of the states of all constituent, entangled particles. Thus, the quantum superstate of n entangled qubits, where each qubit is described by two complex numbers, is described by 2^n complex numbers (Grover, 1996). While doubling the number of bits in a classical computer doubles its processing power, adding qubits results in an exponential growth in computing power. Thus, quantum entanglement, combined with superposition, is a property expected to provide an exponential increase in computing power in quantum computers, ultimately leading to quantum supremacy (Jozsa & Linden, 2003).

3.2 What quantum technologies have solved so far

When in 1994 Peter Shor devised a quantum computer algorithm for integer factorization, it was considered a long-awaited breakthrough in the field of computational complexity (Shor, 1997). Up to that point, the integer factorization problem had no efficient algorithm; this made integer factorization impossible to solve in finite time on any machine (Briggs, 1998). Shor's proposition, which took advantage of quantum effects to reduce the complexity of the solution to integer factorization, was the first practical proof that a quantum algorithm can significantly surpass the capabilities of algorithms running on classical machines. Now, quantum algorithms are widely maintained to bring powerful improvements to combinatorial optimization, differential equation solvers, and simulations of many-body systems (Rønnow et al., 2014). The commercial industries mentioned most frequently include materials and pharmaceuticals, banking and finance, advanced manufacturing, and cybersecurity (Bova et al., 2021). Existing quantum algorithms for the aforementioned applications have been collected and presented by Bharti et al. (2022).

To be practically relevant, quantum algorithms must run on some sort of quantum machine. Nowadays, the largest quantum computers are a 72-qubit Bristlecone

processor created by Google and a 65-qubit Hummingbird by IBM. However, both have a small number of entangled qubits and therefore low efficiency. In addition to increasing the number of qubits, the efficiency of a quantum computer is affected by the quality and functionality of the quantum system at the single-qubit level. Key properties used to assess the quality of quantum systems include *connectivity* (also called system topology), which determines possible entanglement and directly affects the exponential increase in computing power (Hu et al., 2022); metrics of *fidelity*, which assess qubit error rates and are calculated as the ratio of the obtained value to the expected value (see Nielsen, 2002); *relaxation time* (T1), i.e. loss of energy from the system, and *dephasing time* (T2), i.e. the time during which qubit phase stays intact. Some of those metrics are part of the supposed-to-be universal, single-number metric proposed by IBM in 2019 known as *quantum volume* (Cross et al., 2019). Considering the quantum volume, the most powerful quantum computers are believed to be the 32-qubit Quantinuum System Model H2, which achieved a quantum volume of 65,536 (Moses et al., 2023), and the 20-qubit Quantinuum System Model H1-1, which achieved a quantum volume of 524,288 (Morrison, 2023). When the IONQ company announced in 2020 that they had built a 32-qubit computer with an expected quantum volume of 4,000,000, it seemed like a breakthrough in quantum technologies. At that time, Honeywell's 10-qubit System Model H1 was considered the most advanced quantum computer, boasting a quantum volume of 128 (Moore, 2020). However, the stated expected quantum volume value of 4,000,000 is purely speculative and has not yet been confirmed by any scientific publication.

Having discussed the historical significance of Peter Shor's quantum algorithm and the current landscape of quantum computing technology, it becomes evident that the realization of quantum supremacy and the intelligence explosion depends on two factors. The first factor is quantum algorithms that leverage the phenomena of superposition and entanglement to enhance computational capacities beyond the hypercomputational barrier, which is crucial for achieving technological singularity. An example of such a quantum algorithm, vastly superior to known classical solutions, is Shor's integer factorization algorithm. The second basis is quantum machines that execute quantum algorithms. Without a well-functioning machine, even the most powerful algorithm will not have a chance to work, and quantum supremacy cannot be achieved. Nevertheless, the practical feasibility of large-scale quantum computers remains a significant challenge. The next section will explore the issues that limit the power of quantum computers.

3.3 Practical limitations of quantum computers

One of the most important practical issues limiting the computational capabilities of quantum computers is decoherence, i.e., the loss of quantum properties of a system due to interactions with the surrounding environment. Quantum coherence is a vital component of scalable quantum computation (Ladd et al., 2010). To date, the problem of decoherence is an important aspect even in the best modern quantum

computers because the highest reported coherence times regarding quantum computers were T1=113μs and T2=122μs, achieved by the IBM Quantum Falcon (Jurcevic et al., 2021).

Very recent and pioneering studies suggested the possibility of improving qubit coherence even to one hour (for an overview, see Saki et al., 2019); nevertheless, as Ryan-Anderson noticed, it is highly unlikely that a physical qubit will ever be able to reach the precision demanded by large-scale computations. Rather than by reducing decoherence time, the reliable large-scale quantum computations might be achieved by efficient suppression of errors and decoherence with quantum error correction (QEC; see Ryan-Anderson et al., 2021 for an example). QEC usually involves encoding a bit of quantum information into an ensemble of qubits that act together as a single *logical qubit*. However, to truly bring a solution to the decoherence problem, the QEC algorithms must achieve high-level error reduction in a reasonable time using reasonable resources. To date, Google showed that by increasing the number of qubits by 32 they were able to reduce an error rate by 3.76% (0.114% points; Google Quantum, 2023). Most forecasts show that applying QEC to the existing quantum algorithms is expected to require the ensemble of at least 1,000 physical qubits to produce a single logical qubit (Krupansky, 2023).

Because the construction of a quantum computer with a sufficient number of physical qubits for QEC requires a significant breakthrough in quantum technologies, the physical realizability and practical applicability of QEC to reliable large-scale quantum computations are debatable (Krupansky, 2023). To date, there is no promising proposal for a quantum computer design capable of large-scale computations, which is necessary to achieve quantum supremacy and, consequently, superintelligence.

4 Limitations of computability

While there are currently no quantum computers capable of reliably performing large-scale computations, quantum algorithms undoubtedly play a central role in the discussion of quantum supremacy. Achieving quantum supremacy involves quantum algorithms providing effective solutions to problems that were previously deemed unsolvable (Copeland, 2002; Preskill, 2012). The limits of solvability have been established by mathematics; therefore, the discussion of superintelligence and quantum supremacy must be framed within the framework of computability theory.

The history of computability, and therefore formalized algorithms and computers, began in the mid-twentieth century. Kurt Gödel (Gödel & Feferman, 1986), Church, (1936; Turing, (1937) almost simultaneously attempted to define precisely the concept of *computable* expression. These attempts to define what precisely *automatically computable* means resulted in the world-famous Church-Turing Thesis (CTT) that defined the computability of function in terms of a so-called Turing machine: a function on the natural numbers can be computed by an *effective method* if and only if it is computable by the Turing machine (Copeland, 2008). The Turing machine is the most intuitive concept of computability that has been invented; it is a hypothetical state machine that performs a *finite* number of operations on an infinite

one-dimensional tape that acts like the memory of a typical computer. The Turing machine performs the operations according to the provided instruction table that consists of a list of possible states of the machine and possible transitions to other states. At each stage of computation, the Turing machine can read a symbol from a tape cell, change the state based on the read symbol and the instruction table, write a symbol to the cell, or move on (for a broader description of the Turing machine, see De Mol, 2018). Since Turing machines are easy to analyze and capture intuitions about computation as a finite set of operations, it is believed that if a problem can be described in terms of an algorithm, then there must also exist a Turing machine that can execute this algorithm (Davis, 2013). To date, this belief, which reduces the computability of a problem to the existence of a Turing machine for that problem, has not been confirmed nor disproved.

With this formalization of automatically computable functions emerged the first evidence of mathematical problems that cannot be described by the classical Turing machine (Davis, 1965). While introducing the idea of the Turing machine, Alan Turing provided the first formal proof of the *unsolvability* of some problem; he showed that it is impossible to create the general algorithm that decides whether a given program will stop or run forever (i.e., the halting problem; Sipser, 1996). It is now known that there is an infinite number of such undecidable problems. To date, it is impossible to determine whether a Turing machine incorrectly defines computability, or whether there are indeed some problems in our world that can never be solved automatically.

The question about the possibility of an automatic solution to any problem raises yet another important question: are there any natural, physical limits of computation that are related to the finite nature of resources in the real world, such as space and time? The rationale for this question is simple: if computers were infinitely fast and their memory infinitely large, any method would suffice to solve the problem. However, the speed of computers is finite, the memory is finite, and the cost of electricity increases sharply with the consumed computing power. Therefore, different solutions to the same problem can be compared in terms of the resources they use. The aforementioned question led to the development of the field known as *computational complexity* which focuses on classifying problems according to the resources that Turing machines need to solve a problem (Smale, 1997). The amount of resources required depends on the size of the input data; thus, the computational complexity of an algorithm, i.e., the necessary effort to calculate an algorithm, is usually expressed as a function $f(n)$, where f is a function that describes the number of operations to be performed when solving a problem and n is a size of the input data. Consider the following example to get a good understanding of the concept of computational complexity. Sorting an array of digits of length n using the bubble sort algorithm, in the worst possible case when elements of the array are arranged in decreasing order, requires n^2 operations (Astrachan, 2003). Thus, the so-called worst-case time complexity of bubble sort is $f(n) = n^2$; this also means that there exists a Turing machine that can execute the bubble sorting algorithm, and this Turing machine will return the sorted array in time of n^2. The empirical boundary of what is called an effective solution to a problem is a polynomial function. However, many existing problems

can only be described by a Turing machine that has much worse than polynomial computational complexity, e.g., 2^n; this implies an exponential increase in time or storage with the increase in the size of the input data (Hartmanis & Stearns, 1965). Importantly, modern classical computers are almost equivalent to a Turing machine in terms of what they can compute and how fast. Thus, problems that are characterized by exponential computational complexity, e.g., aforementioned integer factorization, cannot be solved efficiently on classical computers (Brent, 2000).

In 1963 Patric Fisher introduced an abstract concept of a non-deterministic Turing machine (Fisher, 1963). Unlike the classical deterministic Turing machine, the non-deterministic Turing machine cannot be realized by any computer because the transition function of the non-deterministic Turing machine returns a set of all possible transitions to other states instead of one deterministic state; thus, a non-deterministic Turing machine can be described in a form of a *tree* of all possible operations and transitions, rather than in a form of a linear sequence of operations which can be performed by a classical computer. In this tree of possible solution paths, all branches are checked simultaneously; this enables non-deterministic Turing machines to have much better computational complexity than classical, deterministic Turing machines, which must perform operations linearly. The introduction of non-deterministic machines does not change the set of Turing-computable, i.e. solvable functions, because the unsolvability of these problems is not related to the number of resources. However, the computational complexity differs for deterministic and non-deterministic Turing machines (Garey & Johnson, 1979).

The field of computational complexity classifies algorithms according to the model of computation (deterministic and non-deterministic Turing machines) and resource demands (i.e., polynomial, exponential). The fundamental complexity class is P (from polynomial time). Problems within this class could be solved with a deterministic Turing machine in polynomial time. Likewise, the NP (from non-deterministic polynomial-time) class contains problems that could be solved with a non-deterministic Turing machine in a polynomial time. Since finding a solution by a non-deterministic Turing machine is proved to be equivalent to a verifying solution by a deterministic Turing machine (see Savitch, 1970), NP problems are usually defined as problems for which classic computers can efficiently check whether the given solution is true. One of the most famous and yet unsolved problems is the question of whether the classes of algorithms calculated with polynomial effort coincide for deterministic and nondeterministic Turing machines (see Baker et al., 1975; Feinstein, 2006). If so, this would imply that a simultaneous search for a solution across all possible solution paths takes the same amount of effort as a linear search across all possible solution paths. The proof of equivalence of P and NP classes would signify a fundamental departure from our current understanding of the world: it implies a world where there is no distinction between parallel and linear operations, and no distinction between solving the problem and merely verifying its correctness. As pointed out by Scott Aaronson, the absence of a distinction between finding a solution and verifying its correctness has profound physical and philosophical consequences, as it equates creation with the simple recognition of creation (Aaronson, 2013b; see also Granade, 2009). Until it is proven whether P = NP or

P$\neq$NP, technological development will remain at a crossroads, awaiting the delineation of further theoretical development paths and the establishment of irrefutable boundaries in computational capabilities.

4.1 Can quantum computers contribute to the emergence of superintelligence?

To synthesize the above discussion, mathematics suggests two distinct computational barriers: the first one limits the solvability of problems and indicates that some problems can never be solved. These are problems for which no Turing machine can be created. The second barrier is the limit of computational complexity. The distinction between P and NP complexity classes suggests that there are some problems that in the physical world have only ineffective solutions. Problems that can be solved only by non-deterministic Turing machines are in principle solvable, but their solution is not physically realizable. Achieving technological singularity through quantum supremacy seems to require crossing both barriers (Copeland, 2002; Goertzel, 2013). However, do these theoretical barriers delineate the actual physical computational barriers?

There are several important open issues regarding the relationship between the CTT, physics, and the possibility of hypercomputation, i.e., computation of unsolvable problems (see Syropoulos, 2008). The CTT itself does not state anything about the world. It does not refer to any epistemological, logical, or physical principles. To extend its influence to the empirical world, many attempts have been made to formulate a physical version of the CTT. One of the first formulations stated that any physical process can be simulated by some Turing machine (see e.g. Deutsch, 1985). This definition was quickly disproved with numerous examples of random and infinite processes in the physical world, for which a Turing machine certainly cannot be created. Piccinini (2011) distinguished between hard CTT formulations, which he called the bold physical CTT, and the modest physical CTT. The bold physical CTT refers to Deutsch's formulation and states that *any physical process* is computable by a Turing machine. The latter one states that *any physically computable function* is Turing-computable. Since the modest physical CTT, like the hard CCT, is falsifiable and open to empirical refutation, but is less stringent, it qualifies as a valid scientific theory and appears to be the most effective tool for evaluating the possibility of hypercomputation proposed so far. If true, the modest physical CTT makes strong claims about the nature of computation, its physical boundaries, and the whole universe because it refutes the possibility of hypercomputation; given the modest physical CTT, the undecidable problems and exponential complexity are inevitably embedded in the world.

The previous decade brought numerous attempts to refute the modest physical CTT; many researchers have proposed various hypothetical physical worlds, techniques, and phenomena that would enable hypercomputation and thus falsify the modest physical CTT (see Németi & Dávid, 2006; Pitowsky, 1990; Shagrir & Pitowsky, 2003). These techniques involve: closed timelike curves (Brun, 2003), general relativistic spacetimes (Hogarth, 1992; Welch, 2006), infinite-precision real numbers (Blum et al., 1989), stuffing an infinite number of steps in a finite time

(Hamkins, 2004) and quantum effects (Kieu, 2002; Ord & Kieu, 2005). However, as Scott Aaronson, a renowned computer scientist and expert in computability problems, shows, all these ideas introduce speculative changes to the laws of physics, or they use only theoretically possible phenomena that cannot be decided whether they belong to our universe (for review, see Aaronson, 2005); thus, rendering these phenomena in our world is highly unrealizable. Scott (and other physicists) point out that there are some basic but non-negligible physical constraints in our universe (Aaronson, 2005; Cotogno, 2003; Sandberg, 1999). One of the most important constraints is the Bekenstein bound; it asserts a particular limit on the amount of information that a finite amount of space can contain (Bekenstein, 1981). This bound implies a limitation of the maximum possible computation rate in finite-sized systems called Bremermann's limit (Bremermann, 1967). Bremermann's limit excludes the possibility of infinite precision or unbounded memory in devices with finite physical dimensions in any place in the whole material universe, even near black holes. In the essay entitled 'The Myth of Hypercomputation', Martin Davis pointed out that hypercomputability claims are usually based, mostly implicitly, on assumptions that involve infinite inputs or resources, e.g., infinite memory, infinitely many steps, infinitely small parts, infinitely fast information transfer, infinitely precise measurement of quantum states (Davis, 2004). These infinities are physically implausible considering modern physics and Bremermann's limit (Müller, 2011; Barrow, 2005). While quantum information theory theoretically allows a single qubit to store an infinite amount of information, as discussed in the third Section, these theoretical possibilities are not realizable in the physical world (Ziegler, 2005). Therefore, there is no premise for the claim that superposition is a phenomenon that can be used to achieve quantum supremacy and intelligence explosion.

Contrary to popular belief that unsolvability is a kind of anomaly in the universe and humanity should seek a solution to this problem, some scientists consider unsolvability as a kind of elemental randomness in mathematics (for more details, see Chaitin, 2002a, 2002b), and thus as an inherent element of reality. In his paper on physics simulations, Feynman (1982) argued that the barrier of uncomputability is insurmountable and should be widely accepted as such in the scientific community (Bernstein & Vazirani, 1997). Physics suggests that since crossing the Turing barrier involves infinity, this barrier may indeed exist (Barrow, 2005). If the Turing computability barrier were seen as a principle of physics, then the world would be as expected to be: with limits of light speed (Cockshott et al., 2008), laws of thermodynamics (Sandberg, 1999), Heisenberg's uncertainty principle (Cockshott et al., 2008), finite measurement precision (Deutsch, 1997), and no back-time travels (Aaronson & Watrous, 2009); in other words, without infinity. This physical limitation must be obeyed equally by classical and quantum computational machines, at least until a breakthrough in physics. Until then, neither classical, nor quantum computers can perform hypercomputation. Quantum computers do not fundamentally change the discussion about the emergence of superintelligence through technological singularity.

Quantum computers carried yet another promise: they would speed up computing enough to efficiently solve problems that do not have efficient solutions. As mentioned in the third Section, by means of superposition and entanglement qubits

were thought to represent the 2^n amount of information. This property of qubits was thought to allow humanity to solve NP problems in polynomial time (see Aaronson, 2007). However, the crucial issue of computational techniques is not about the amount of information that could be stored in a given moment of computation; it is rather about the amount of information that could be received, decoded, and measured to establish the result. In 1973, Alexander Holevo published a theorem that established an upper bound to the amount of information that can be known about a quantum state. Holevo proved that at most one bit of information can be extracted from one qubit; thus, although purely theoretically a quantum state of n qubits can carry 2^n amount of information, such a state can return at most n bits of decodable information. Thus, to date, there is no premise that quantum algorithms will effectively solve *all* NP problems. Why, then, have quantum algorithms brought about such big computational improvements for many computationally demanding problems? Taking advantage of quantum effects, quantum algorithms support the simulation of systems where quantum phenomena play an important role (Rønnow et al., 2014); for quantum problems, they can provide more efficient (quantum) solutions compared to classical algorithms. To account for the possibilities of quantum algorithms, which differ from the classical ones, computational complexity theory distinguished a class for problems that could be efficiently solved by quantum computers called BQP (for bounded error, quantum, polynomial time; Nielsen & Chuang, 2010). The relationship of BQP to NP is not known, though it is conjectured that NP $\not\subset$ BQP and quantum algorithms could not solve at least the hardest and thus most important problems of the NP class, i.e., NP-complete problems, in polynomial time (Aaronson et al., 2022; Watorous, 2008; Knill & Nielsen, 2000; Aaronson, 2005). Therefore, quantum algorithms have the potential to significantly enhance solutions for quantum-related problems. However, there is no evidence to suggest that focusing solely on solving quantum-related problems would necessarily lead to technological singularity.

In conclusion, complexity classes are influenced by quantum effects and differ from those considered only within classical physics; the computability limits, however, remain the same for both quantum and classical cases. If a quantum computer can perform some operation, a classical Turing machine can simulate a quantum computer performing this operation. Quantum computers do not exponentially accelerate computing power in general and do not lead to technological singularity, as they might bring exponential speed-up only in special quantum cases. Given the presented discussion on hypercomputation and computational complexity theory, there is insufficient evidence to support the claim that quantum technologies could alter the computational barrier and contribute to the emergence of superintelligence through technological singularity.

5 The dream of an artificial mind

As outlined in the previous sections, quantum supremacy, the theoretical point at which a quantum computer can perform tasks beyond the capabilities of classical computers, remains a topic of debate and exploration. While there have been

significant advancements in quantum computing, there is still vast uncertainty surrounding whether quantum supremacy is achievable, particularly given the complex challenges involved in building and scaling quantum systems and limits of computability. Current discoveries in modern physics do not provide a solid foundation for asserting the feasibility of quantum supremacy. Thus, achieving *quality* or *speed* superintelligence by means of quantum technological singularity and intelligence explosion (e.g., Bostrom, 1998; Goertzel, 2013; Zheng & Akhmad, 2017; Miller, 2019) does not seem possible. However, because quantum computers can bring significant acceleration to quantum-defined problems, some researchers continue to link the development of quantum technology to the development of superintelligence (e.g. Zheng & Akhmad, 2017; Yampolskiy, 2018). In the following sections, we will discuss two alternative theses that connect quantum technologies with the emergence of superintelligence but do not involve a technological singularity. The first thesis claims that modern AI algorithms will indeed benefit significantly from the development of quantum technologies. The implicit assumption of this approach is that general intelligence is an algorithmically definable problem, it can be solved by mere intensification of one already known computing strategy, and it can be better defined in a quantum way than in a classical way. The second thesis denies that intelligence can be described by any algorithm; however, it claims that the human mind is *based* on quantum effects. Therefore, quantum computers may be better suited to simulate the mind than classical computers.

5.1 The quantum implementation of the mind

Current AI models are near-perfectly functioning specialized systems that support people and expand their skills in many areas. It seems that there is no such specialized task in which one of the AI systems does not exceed human abilities, and quantum computing could provide significant speed-ups, expanding the range of possibilities of artificial systems (Liu et al., 2021). In fact, with a sufficiently large number of engineers, time, and computing resources, it is possible to create millions of specialized software programs for millions of specific situations and problems. These millions of programs will be functionally equivalent in some way to the human mind. This approach to general intelligence is reflected in Marvin Minsky's The Society of Mind (Minsky, 1988), which describes human intelligence as a set of simple specialized processes. Pierro Scaruffi calls this solution brute-force AI (Scaruffi, 2018). Brute-force AI, although achievable, does not arouse much enthusiasm in researchers today. Despite the impressive results, it remains only a complex sequence of mathematical operations performed on a powerful computing machine, which is difficult to call intelligence regardless of the adopted definition. There is ample evidence that intelligence is more than a sum of specific task-solving subcomponents (Landgrebe & Smith,2022). Although this approach could solve particularly important problems, it lacks the ability of general reasoning in an uncertain environment, which is usually referred to as the most important aspect of any general intelligent agent (see Bostrom, 2014). Thus, even a significant acceleration of existing

specialized AI algorithms may only yield limited progress in the field of superintelligence. The algorithms that might lead to AGI and superintelligence should computationally define the ability of general reasoning; for quantum computers to be essential in implementing such algorithms, they must offer advantages over classical computing in terms of defining and executing these algorithms. The following sections focus on the two primary approaches to computationally modeled reasoning and assess the potential enhancements of quantum technologies.

Today, one of the most promising approaches to the implementation of artificial general reasoning is reinforcement learning (RL), i.e., active learning from interactions with a complex and uncertain environment (Sutton & Barto, 2018). RL consists of learning via trial and error through a process designed to maximize the reward returned from the environment. Among various approaches to the implementation of artificial general reasoning, RL most closely resembles the form in which humans acquire skills and knowledge and thus seems to be the best approach to achieve speed or quality superintelligence. With the possibility of building any precise and specified environment, RL is recognized as one of the most universal and powerful approaches to the implementation of active learning of artificial agents (Nian et al., 2020). If RL algorithms could be effectively represented in quantum form, this would be an important premise in considering the potential contribution of quantum algorithms to the emergence of superintelligence. Indeed, it has been shown that, under certain assumptions, quantum effects can significantly improve the RL process (for review, see Dunjko et al., 2017). Various studies have shown a quadratic speed-up in the agent's decision-making process (Agunbiade, 2022; Albarran-Arriagada et al., 2020; Sriarunothai et al., 2018; Saggio et al., 2021). Although these achievements are certainly significant, Dunjko and colleagues pointed out that they do not appear to have contributed to a breakthrough in the RL. The foundation of RL lies in the agent's interactions with the environment; the quality of the model and learning efficiency are directly related to the number of interactions. This creates a quantization bottleneck that prevents quantum RL from a ground-breaking improvement and is often considered an inherent limitation of the entire RL paradigm (Dunjko et al., 2017). Thus, as mentioned at the beginning of this section, the expectation of breakthrough advances in RL brought by quantum computers should be related to the fact that reinforcement learning is a problem that is better described by a quantum system than a classical one, rather than to the quantum acceleration of computation in the RL paradigm. Recently, Li et al. (2020) suggested that RL indeed might be better described by a quantum system. They demonstrated higher performance of a decision-making simulation in quantum than in a classical RL environment, which they interpreted as evidence of the link between neuroscience data and quantum mechanics. However, this superiority of quantum models over classical ones does not necessarily indicate that the human brain is based on quantum effects. Some classical processes have been shown to generate quantum properties under certain conditions; it is not enough to demonstrate the better performance of a quantum algorithm over a classical one under specific conditions to prove that a process is quantum in nature (Ivakhnenko et al., 2018). To date, there is no sufficient evidence that RL is inherently quantum-based, and consequently, there is limited

evidence suggesting that quantum technologies could enhance RL to the point of achieving general intelligence level capabilities.

Another frequently mentioned approach to computational modeling of general intelligence involves probabilistic graphical models, especially Bayesian models (van de Schoot et al., 2017). Bayesian models determine the probability of a certain event based on prior knowledge and beliefs about events and conditions with which it might be associated (Bayes, 1763). Since quantum computing is inherently probabilistic, this property could be used to significantly increase computational capacity in modeling Bayesian solutions (Benedetti et al., 2021). It is believed that the probability distribution of classical data could be directly and naturally represented using a pure quantum state instead of complex analytical equations (e.g. Cheng et al., 2018). Nonetheless, despite the superiority of the quantum data representation, inference in Bayesian networks is an NP-hard problem (due to their graphical representation) and even heuristic algorithms involve exponential complexity in the number of nodes (Kwisthout & Rooij, 2013). The foundation of Bayesian inference is belief propagation, which requires the conversion of a network into a structure named the 'junction tree', regardless of whether the network is classical or quantum. Finding an optimal junction tree is an NP-hard problem, and the complexity of the algorithms scales as the width of the tree increases. As noted in the fourth Section, the BQP class that distinguishes problems efficiently solvable by quantum algorithms does not overlap with the NP class, and quantum algorithms certainly cannot solve the hardest problems from NP (Aaronson et al., 2022). Besides problems with computational complexity, researchers usually mention yet another general problem with the Bayesian approach to modeling general reasoning (for an overview, see Bowers & Davis, 2012; but see also Griffiths et al., 2012 for a reply). Classical Bayesian models try to estimate what people rationally should choose, rather than what they actually chose; biological, evolutionary, and psychological findings suggest that the human mind was designed to find *satisfactory*, rather than optimal solutions; these solutions are adaptive to many conditions, but never optimal (see Gigerenzer & Brighton, 2009; Kahneman & Tversky, 1996). Recently, Trueblood et al. (2017) and Wichert et al. (2020) attempted to address the problem of paradoxical and irrational human inference using quantum probability theory. Wichert et al. showed a relationship between quantum probability waves and empirical findings from famous probability problems (e.g. Prisoner's Dilemma, Two-Stage Gambling Game). However exciting, these results should be interpreted with caution, as they have not been tested in real-world complex decision problems and their relevance to general reasoning and intelligence is modest.

The computational approaches to the general reasoning shown above hold an important implicit assumption: general reasoning and intelligence are Turing-computable, i.e., solvable problems; further, to be physically modeled, this problem can be solved effectively. As mentioned in the second Section, AGI is expected to replicate human abilities other than general reasoning, such as emotional, moral, or social intelligence. Contrary to analytical skills, these abilities are largely not subject to learning, as is classically understood, being more innate than acquired in nature. Thus, the possibility of creating an artificial agent at least as good as a human in any field, with the help of AI algorithms, requires adopting one of the strongest

physicalist theories of the mind, i.e., the computational theory of mind, to make emotional, moral, or social abilities not only physical processes but also Turing-computable processes (McCulloch & Pitts, 1943; Putnam, 1967; Fodor, 1975). According to these theories, phenomena both at the macro and micro levels, such as sociological and psychological phenomena, respectively are subject to certain rules and laws and are reducible to physics. For example, in the view of reductive physicalism (Kim, 1998), it is plausible to argue that not only do such laws exist, but also that appropriate bridge laws will certainly be constructed that will link them with the laws of physics, and that this is only a matter of time and technological development. Similarly, the adoption of eliminative materialism (Churchland, 1981) allows for the assumption that the artificial agent may manifest emotional and social abilities. In this approach, there is nothing that is not a brain process. As the brain is a physical object and the pro-spaces governing it are subject to the laws of physics, it is acceptable to say that emotional and social processes are simulable in the universe and that this also is only a matter of time.

Although intuitive, assumptions of computability have been widely criticized by many researchers (e.g., Roitblat, 2020; Landgrebe & Smith, 2022; for a discussion, see Tallis and Alexander, 2008). Back in the 1960s, Gödel's incompleteness theorem was used to demonstrate some strict limitations on the mechanization of thought, understanding, consciousness, and awareness (Lucas, 1961; Wang, 2016). This issue was discussed most thoroughly by Roger Penrose; Penrose developed a Godelian argument to show that no Turing machine or algorithm can in principle simulate the self-awareness aspect of thought essential to the process of conscious understanding as experienced by human minds (Penrose, 1989). Thus, although there is some consensus on the physicality of mental processes, their Turing-computability is yet not proven; according to the modest physical CTT, not every physical process that exists in the world must be solvable. The problem of finding an *effective* solution for AGI is usually not mentioned at all (Šekrst & Skansi, 2021).

In the 1990s, David Wolpert developed the famous 'No Free Lunch' (NFL) theorems for search problems, machine learning and optimization (Wolpert & Macready, 1995, 1997). NFL theorems demonstrate that no learning algorithm could perform well in all classes of problems, which means that no algorithm could excel at everything. There is a heated debate among the scientific community about the impact of the NFL on the possibility of creating general intelligence or general reasoning systems. NFL theories highlight the obstacles towards the emergence of the AGI by either classical or quantum approaches. There are no free lunches, however, the problem of whether superintelligence itself is a free lunch remains unsolved and quantum technologies do not seem to bring a new quality to this discussion. Quantum technologies neither guarantee the computability of general reasoning, nor are there sufficient premises to claim that, if general reasoning is computable, it is better expressed in a quantum way to effectively use quantum technologies for its development (Aaronson, 2013a).

5.2 The quantum emergence of the mind

The second proposal that links quantum technologies to the development of super-intelligence presents an opposite point of view compared to the computational approach to reasoning discussed earlier. It denies that intelligence can be described by any algorithm; instead, it reaches into the nature of the mind, claiming that it arises from quantum effects in the brain.

Over the past few decades, a new field within AI called artificial life has developed. It is centered around the idea of mimicking biological behaviors in non-living systems under controlled laboratory conditions (Bedau, 2007). In his paper, Chaitin (2010) asked the question that has haunted mankind for decades: is it possible to mathematically prove evolution? Following Chaitin, another question may be carefully asked: Is the quantum world, or the classical world, the better environment for evolution? That is: does evolution better describe the phenomena of classical or quantum physics? Martin-Delgado (2012) tried to answer this question by creating a model that takes into account the impact of quantum effects on evolution. Based on its results, he noticed that the quantum environment may be more natural to drive evolution since it is inextricably burdened with error, similar to Nature. If so, quantum technologies can be considered essential for the emergence of superintelligence.

These preliminary results draw attention to the concept of quantum biomimetics (Alvarez-Rodriguez et al., 2014). Biomimetics studies the functioning of living things that serve as models in various applications through reverse engineering. The first successful attempts to simulate quantum artificial life, i.e., to model basic features of living such as reproduction, mutation, and interaction, have already taken place (Alvarez-Rodriguez et al., 2016). In 2018, to support simulation with empirical results, Alvarez-Rodriguez and colleagues demonstrated the first experimental realization of quantum life in an IBM quantum computer (Alvarez-Rodriquez, 2018). These results may support the whole-brain emulation hypothesis that the realization of AGI could be based on the natural recreation of artificial individuals in quantum environments, provided that the quantum environment, rather than the classical one, would favor mind processes.

Such assumptions guided Roger Penrose and Stuart Hameroff's hypothesis about the quantum origin of consciousness, called Orchestrated Objective Reduction (Orch OR; Hameroff & Penrose, 1996, 2014). Penrose and Hameroff claimed that consciousness is based on non-computable quantum processes inside neurons, specifically within microtubules, which would arise from a yet-to-be-discovered quantum theory of gravity. However, many scientists argue that the Orch OR hypothesis lacks a solid theoretical basis and there is little or even no evidence supporting the idea that quantum effects play any significant role in emerging consciousness and the human mind. Critics of Orch OR note that the warm and wet environment of the brain does not seem conducive to sustain quantum phenomena; decoherence would occur so rapidly that quantum phenomena would not have any real impact on humans (see Tegmark, 2007).

The second argument against the quantum emergence of mind is based on empirical observations: it does not seem that humans work in the same way as quantum computers – they do not cope better with tasks where quantum computers yield

better results than classical ones, such as factorization (Aaronson, 2013a). Problems considered as BQP do not seem crucial for the evolution and survival of humans. Until there is empirical or theoretical evidence that mental processes are based on quantum effects, claims that quantum computers can be used to create an artificial mind, and thus superintelligence, lack justification.

6 Conclusions

In the presented work, we discussed how quantum computers could contribute to the emergence of superintelligence. The first and most popular scenario involves the emergence of superintelligence through quantum supremacy, which is caused by the acceleration of computation beyond the Turing barrier using quantum computers. We framed this discussion within the context of the possibility of hyper-computation and we showed that quantum computers cannot accelerate many key problems; the computational barriers are similar for classical and quantum computers. Thus, according to the perspective of modern physics, achieving quantum supremacy does not appear feasible. The second scenario involves the creation of AGI through the implementation of quantum versions of existing AI algorithms, which would significantly expand the capabilities of these algorithms. We demonstrated that there is no evidence supporting the superiority of the quantum-based definition of these algorithms. Furthermore, we showed that this scenario implicitly assumes the Turing-computability of the mind, which is also a subject of heated debate. The third scenario involves the creation of AGI through the quantum emergence of the mind. We demonstrated that, thus far, there is no evidence supporting the idea that mental processes are based on quantum effects. Therefore, there is no premise that quantum technologies can contribute to the emergence of superintelligence in this scenario. As it stands, quantum technologies do not seem to bring the expected breakthrough to the problem of building an artificial, truly intelligent agent.

Although progress is being made in many areas and human natural abilities are daily supported by artificial systems to such an extent that the term 'cyborg' can be applied to any human being, superintelligence far exceeds the current scope of technology, scientific knowledge, definitions, and comprehension. An endlessly long discussion is required to define what the phrase 'all cognitive tasks' included in Bostrom's definition means and to determine how to achieve a system that will reach the state humans have attained after thousands of years of evolution, without evolution. It does not appear that quantum technologies can contribute significant value to this discussion. This does not mean that superintelligence objectively cannot exist. According to the concept of cognitive closure popularized by Colin McGinn in 'The Problem of Consciousness' (McGinn, 1991), every cognitive system has its own boundaries of cognition and cannot know beyond these boundaries. Within this view, it is possible to imagine cognitive systems possessing cognitive boundaries broader than humans. On the other hand, David Deutsch, in 'The Beginning of Infinity' (Deutsch, 2011), presented an interesting physicalist point of view. He noticed that, if the universe is driven by universal

laws and the theory of everything is possible, then there is nothing in the universe that humanity cannot understand. Universal laws of nature consist of physicalist formulas and rules that might therefore be discovered and understood, meaning that human beings are the highest form of intelligence that could exist. While it is reasonable to expect quantum computers to bring artificial systems closer to human cognitive limits and enhance natural human abilities, they do not seem to open the way towards transgressing those limits to achieve quality superintelligence, at least not with our current understanding of the laws governing the universe.

Declarations

Competing interests Authors declare no competing interests

References

Aaronson, S. (2005). Guest Column: NP-complete problems and physical reality. *ACM SIGACT News, 36*(1), 30–52. https://doi.org/10.1145/1052796.1052804

Aaronson, S. (2007). The Orion Quantum Computer Anti-Hype FAQ. *Shtetl-Optimized*. Retrieved May 11, 2023, from https://scottaaronson.blog/?p=198

Aaronson, S. (2010). BQP and the polynomial hierarchy. *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, pp. 141–150. https://doi.org/10.1145/1806689.1806711

Aaronson, S. (2013a). Penrose. In *Quantum computing since Democritus* (pp. 150–159). Cambridge University Press. https://doi.org/10.1017/CBO9780511979309.012

Aaronson, S. (2013b). Why philosophers should care about computational complexity. *Computability: Turing Gödel Church and Beyond,261*, 327.

Aaronson, S., & Watrous, J. (2009). Closed timelike curves make quantum and classical computing equivalent. *Proceedings of the Royal Society A: Mathematical Physical and Engineering Sciences, 465*(2102), 631–647. https://doi.org/10.1098/rspa.2008.0350

Aaronson, S., Ingram, D., & Kretschmer, W. (2022). The Acrobatics of BQP. In S. Lovett (Ed.), *37th Computational Complexity Conference (CCC 2022)* (Vol. 234, p. 20:1-20:17). Schloss Dagstuhl – Leibniz-Zentrum für Informatik. https://doi.org/10.4230/LIPIcs.CCC.2022.20

Agunbiade, A. (2022). *Quantum Computing & Reinforcement Learning: Partners in Achieving Artificial General Intelligence*. SSRN. https://doi.org/10.2139/ssrn.4197291

Albarrán-Arriagada, F., Retamal, J. C., Solano, E., & Lamata, L. (2020). Reinforcement learning for semi-autonomous approximate quantum eigensolver. *Machine Learning: Science and Technology, 1*(1), 015002. https://doi.org/10.1088/2632-2153/ab43b4

Alvarez-Rodriguez, U., Sanz, M., Lamata, L., & Solano, E. (2014). Biomimetic cloning of quantum observables. *Scientific Reports, 4*(1), 1–4. https://doi.org/10.1038/srep04910

Alvarez-Rodriguez, U., Sanz, M., Lamata, L., & Solano, E. (2016). Artificial life in quantum technologies. *Scientific Reports, 6*(1), 1–9. https://doi.org/10.1038/srep20956

Alvarez-Rodriguez, U., Sanz, M., Lamata, L., & Solano, E. (2018). Quantum artificial life in an IBM quantum computer. *Scientific Reports, 8*(1), 1–9. https://doi.org/10.1038/s41598-018-33125-3

Astrachan, O. (2003). Bubble sort: An archaeological algorithmic analysis. *ACM Sigcse Bulletin, 35*(1), 1–5.

Avaliani, A. (2004). *Quantum computers (arXiv:cs/0405004)*. arXiv. https://doi.org/10.48550/arXiv.cs/0405004

Baaquie, B. E., & Kwek, L. C. (2023). *Quantum computers: Theory and algorithms*. Springer Nature.

Baker, T., Gill, J., & Solovay, R. (1975). Relativizations of the P=?NP question. *SIAM Journal on Computing, 4*(4), 431–442.

Barrow, J. D. (2005). *The infinite book: A short guide to the boundless, timeless and endless*. Jonathan Cape.

Bayes, T. (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. *Philosophical Transactions of the Royal Society of London, 53*, 370–418.

Bedau, M. A. (2007). Artificial life. In M. Matthen & C. Stephens (Eds.), *Philosophy of biology* (pp. 585–603). North-Holland. https://doi.org/10.1016/B978-044451543-8/50027-7

Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D, 23*, 287–298. https://doi.org/10.1103/PhysRevD.23.287

Benedetti, M., Coyle, B., Fiorentini, M., Lubasch, M., & Rosenkranz, M. (2021). Variational inference with a quantum computer. *Physical Review Applied*, *16*(4), 044057. https://doi.org/10.1103/PhysRevApplied.16.044057

Benenti, G., Casati, G., & Strini, G. (2004). *Principles of quantum computation and information-volume I: Basic concepts*. World scientific.

Benioff, P. (1982). Quantum mechanical models of Turing machines that dissipate no energy. *Physical Review Letters, 48*(23), 1581–1585. https://doi.org/10.1103/PhysRevLett.48.1581

Bernstein, E., & Vazirani, U. (1997). Quantum complexity theory. *SIAM Journal on Computing, 26*(5), 1411–1473. https://doi.org/10.1137/S0097539796300921

Bharti, K., Cervera-Lierta, A., Kyaw, T. H., Haug, T., Alperin-Lea, S., Anand, A., Degroote, M., Heimonen, H., Kottmann, J. S., Menke, T., Mok, W.-K., Sim, S., Kwek, L.-C., & Aspuru-Guzik, A. (2022). Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, *94*(1), 015004. https://doi.org/10.1103/RevModPhys.94.015004

Bloch, F. (1946). Nuclear induction. *Physical Review*, *70*(7–8), 460–474. https://doi.org/10.1103/PhysRev.70.460

Blum, L., Schub, M., & Smale, S. (1989). On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *American Mathematical Society, 21*(1), 1–46. https://doi.org/10.1090/S0273-0979-1989-15750-9

Born, M., & Wiener, N. (1926). A new formulation of the laws of quantization of periodic and aperiodic phenomena. *Journal of Mathematics and Physics, 5*(1–4), 84–98. https://doi.org/10.1002/sapm19265184

Bostrom, N. (1998). How long before superintelligence? *International Journal of Futures Studies*, 2.

Bostrom, N. (2013). Existential risk prevention as global priority. *Global Policy, 4*(1), 15–31.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin, 138*(3), 389.

Bova, F., Goldfarb, A., & Melko, R. (2021). Commercial applications of quantum computing. *EPJ Quantum Technology, 8*(1), 2. https://doi.org/10.1140/epjqt/s40507-021-00091-1

Bremermann, H. J. (1967). Quantum noise and information. *Proceedings of the fifth berkeley symposium on mathematical statistics and probability* (Vol. 4, pp. 15–20).

Brent, R. P. (2000). Recent progress and prospects for integer factorisation algorithms. In D.-Z. Du, P. Eades, V. Estivill-Castro, X. Lin, & A. Sharma (Eds.), *Computing and combinatorics* (pp. 3–22). Springer Berlin Heidelberg.

Briggs, M. E. (1998). *An introduction to the general number field sieve* [Master's thesis, Virginia Tech]. http://hdl.handle.net/10919/36618

Brun, T. (2003). Computers with closed timelike curves can solve hard problems efficiently. *Foundations of Physics Letters, 16*(3), 245–253. https://doi.org/10.1023/A:1025967225931

Callaghan, V., Miller, J., Yampolskiy, R., & Armstrong, S. (2017). *Technological singularity*. Springer.

Chaitin, G. J. (2002a). *Exploring randomness*. Springer.

Chaitin, G. J. (2002b). Paradoxes of randomness and the limitations of mathematical reasoning. *Complexity,7*(5), 14–21. https://doi.org/10.1002/cplx.10038

Chaitin, G. J. (2010). *To a mathematical theory of evolution and biological creativity*. Department of Computer Science, The University of Auckland.

Chalmers, D. J. (2016). The singularity: A philosophical analysis. In S. Schneider (Ed.), *Science fiction and philosophy* (1st ed., pp. 171–224). Wiley. https://doi.org/10.1002/9781118922590.ch16

Cheng, S., Chen, J., & Wang, L. (2018). Information perspective to probabilistic modeling: Boltzmann machines versus born machines. *Entropy, 20*(8), 583. https://doi.org/10.3390/e20080583

Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics, 58*(2), 345–363. https://doi.org/10.2307/2371045

Churchland, P. M. (1981). Eliminative materialism and propositional attitudes. *The Journal of Philosophy, 78*(2), 67–90.

Cockshott, P., Mackenzie, L., & Michaelson, G. (2008). Physical constraints on hypercomputation. *Theoretical Computer Science, 394*(3), 159–174. https://doi.org/10.1016/j.tcs.2007.12.009

Copeland, B. J. (2002). Hypercomputation. *Minds and Machines*, *12*, 461–502. https://doi.org/10.1023/A:1021105915386

Copeland, B. J. (2004). Hypercomputation: Philosophical issues. *Theoretical Computer Science,317*(1–3), 251–267.

Copeland, B. J. (2008). The Church-Turing thesis. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. The Metaphysics Research Lab, Stanford University.

Copeland, B. J., & Shagrir, O. (2011). Do accelerating Turing machines compute the uncomputable? *Minds and Machines, 21*(2), 221–239. https://doi.org/10.1007/s11023-011-9238-y

Cotogno, P. (2003). Hypercomputation and the Physical Church-Turing thesis. *The British Journal for the Philosophy of Science, 54*(2), 181–223.

Cotogno, P. (2009). A brief critique of pure hypercomputation. *Minds and Machines, 19*(3), 391–405.

Cross, A. W., Bishop, L. S., Sheldon, S., Nation, P. D., & Gambetta, J. M. (2019). Validating quantum computers using randomized model circuits. *Physical Review A, 100*(3), 032328. https://doi.org/10.1103/physreva.100.032328

Davis, M. (1965). *The undecidable: Basic papers on undecidable propositions, unsolvable problems and computable functions*. Dover Publication.

Davis, M. (2004). The myth of hypercomputation. In C. Teuscher (Ed.), *Alan Turing: Life and legacy of a great thinker* (pp. 196-211). Springer-Verlag.

Davis, M. (2013). *Computability and unsolvability*. Courier Corporation.

De Mol, L. (2018). *Turing Machines*. Retrieved April 11, 2023, from http://seop.illc.uva.nl/entries/turing-machine/

Deutsch, D. (1985). Quantum theory, the church-turing principle and the universal quantum computer. *Proceedings of the Royal Society of London Series A Mathematical and Physical Sciences, 400*(1818), 97–117.

Deutsch, D. (1997). *The Fabric of Reality*. Penguin.

Deutsch, D. (2011). *The beginning of infinity: Explanations that transform the world*. Viking Adult.

Dunjko, V., Taylor, J. M., & Briegel, H. J. (2017). Advances in quantum reinforcement learning. *2017 IEEE International Conference on Systems, Man, and* Cybernetics *(SMC)*. https://doi.org/10.1109/smc.2017.8122616

Feinstein, C. A. (2006). *An elegant argument that P is not NP (arXiv:cs/0607093)*. arXiv. https://doi.org/10.48550/arXiv.cs/0607093

Feynman, R. (1982). Simulating physics with computers. *International Journal of Theoretical Physics*, *21*, 467–488. https://doi.org/10.1007/BF02650179

Fischer, P. C. (1963). On computability by certain classes of restricted Turing machines. In *Proceedings of the Fourth Annual Symposium on Switching Circuit Theory and Logical Design (swct 1963)* (pp. 23–32). IEEE.

Flasiński, M. (2016). *Introduction to artificial intelligence*. Springer.

Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard University Press.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-Completeness*. Freeman.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107–143.

Gödel, K., & Feferman, S. (1986). *Kurt Gödel: Collected works: Volume III: Unpublished essays and lectures* (Vol. 3). Oxford University Press on Demand.

Goertzel, B. (2013). Artificial general intelligence and the future of humanity. In M. More & N. Vita-More (Eds.), *The transhumanist reader* (pp. 128–137). Wiley.

Goertzel, B., & Ikle, M. (2012). Mind uploading (introduction to a special issue on this topic). *International Journal of Machine Consciousness,4*, 1–3.

Good, I. J. (1966). Speculations concerning the first ultraintelligent machine. In F. L. Alt & M. Rubinoff (Eds.), *Advances in computers* (Vol. 6, pp. 31–88). Elsevier. https://doi.org/10.1016/S0065-2458(08)60418-0

Google Quantum, A. I. (2023). Suppressing quantum errors by scaling a surface code logical qubit. *Nature,614*, 676–681. https://doi.org/10.1038/s41586-022-05434-1

Granade, C. (2009). *Why complexity matters: A brief tour* [PDF]. https://www.cgranade.com/research/papers/why-complexity-matters.pdf

Griffiths, T. L., Chater, N., Norris, D., & Pouget, A. (2012). How the Bayesians got their beliefs (and what those beliefs actually are): Comment on Bowers and Davis (2012). *Psychological Bulletin, 138*(3), 415–422. https://doi.org/10.1037/a0026884

Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (pp. 212–219). https://doi.org/10.1145/237814.237866

Gunia, A., & Indurkhya, B. (2017). A prototype to study cognitive and aesthetic aspects of mixed reality technologies. In *2017 3rd IEEE International Conference on Cybernetics (CYBCONF)* (pp. 1–6). IEEE.

Gunia, A., Sołtysik, M., & Jarosz, S. (2024). Robot ethics and artificial morality. *Artificial intelligence, management and trust* (pp. 127–143). Routledge.

Hagar, A., & Korolev, A. (2006). Quantum hypercomputability? *Minds and Machines, 16*, 87–93.

Hagar, A., & Korolev, A. (2007). Quantum hypercomputation—hype or computation? *Philosophy of Science, 74*(3), 347–363.

Hameroff, S., & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and Computers in Simulation,40*(3–4), 453–480.

Hameroff, S., & Penrose, R. (2014). Consciousness in the universe: A review of the 'orch or'theory. *Physics of life Reviews,11*(1), 39–78.

Hamkins, J. D. (2004). *Supertask computation. Classical and new paradigms of computation and their complexity hierarchies* (pp. 141–158). Springer.

Hartmanis, J., & Stearns, R. E. (1965). On the computational complexity of algorithms. *Transactions of the American Mathematical Society, 117*, 285–306.

Hauskeller, M. (2014). *Better humans? Understanding the enhancement project*. Routledge.

Herrnstein, R. J., & Murray, C. (2010). *The bell curve: Intelligence and class structure in American life*. Simon and Schuster.

Hodges, A. (2005). *Can quantum computing solve classically unsolvable problems? (arXiv:quant-ph/0512248)*. arXiv. https://doi.org/10.48550/arXiv.quant-ph/0512248

Hogarth, M. (1992). Does general relativity allow an observer to view an eternity in a finite time? *Foundations of Physics Letters, 5*, 173–181.

Holevo, A. S. (1973). Statistical problems in quantum physics. *Proceedings of the Second Japan-Ussr Symposium on Probability Theory* (pp. 104–119).

Hu, W., Yang, Y., Xia, W., Pi, J., Huang, E., Zhang, X. D., & Xu, H. (2022). Performance of superconducting quantum computing chips under different architecture design. *Quantum Information Processing*, *21*(7), 237. https://doi.org/10.1007/s11128-022-03571-0

Indurkhya, B. (2019). Is morality the last frontier for machines? *New Ideas in Psychology, 54*, 107–111.

Ivakhnenko, O. V., Shevchenko, S. N., & Nori, F. (2018). Simulating quantum dynamical phenomena using classical oscillators: Landau-Zener-Stückelberg-Majorana interferometry, latching modulation, and motional averaging. *Scientific Reports, 8*(1), 1–11. https://doi.org/10.1038/s41598-018-28993-8

Jozsa, R., & Linden, N. (2003). On the role of entanglement in quantum-computational speed-up. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 459*(2036), 2011–2032.

Jurcevic, P., Javadi-Abhari, A., Bishop, L. S., Lauer, I., Bogorin, D. F., Brink, M.,. .. & Gambetta, J. M. (2021). Demonstration of quantum volume 64 on a superconducting quantum computing system. *Quantum Science and Technology, 6*(2), 025020.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*(3), 582–591. https://doi.org/10.1037/0033-295X.103.3.582

Kieu, T. (2002). Quantum hypercomputation. *Minds and Machines, 12*(4), 541–561. https://doi.org/10.1023/A:1021130831101

Kim, J. (1998). *Mind in a physical world: An essay on the mind–body problem and mental causation*. MIT Press.

Kipke, R. (2013). What is cognitive enhancement and is it justified to point out this kind of enhancement within the ethical discussion? In *Cognitive enhancement* (pp. 145–157). Springer.

Knill, E. H., & Nielsen, M. A. (2000). *Theory of quantum computation (arXiv:quant-ph/0010057)*. arXiv. https://doi.org/10.48550/arXiv.quant-ph/0010057

Koene, R. A. (2013). Uploading to Substrate-Independent Minds. *The transhumanist reader: Classical and contemporary essays on the science, technology, and philosophy of the human future*, pp. 146–156.

Krupansky, J. (2023). Why I'm Rapidly Losing Faith in the Prospects for Quantum Error Correction. *Medium*. Retrieved May, 18, 2024, from https://jackkrupansky.medium.com/why-im-rapidly-losing-faith-in-the-prospects-for-quantum-error-correction-5aed92ee7a52

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin.

Kutt, K., Gunia, A., & Nalepa, G. J. (2015). Cognitive enhancement: How to increase chance of survival in the jungle? In *2015 IEEE 2nd International Conference on Cybernetics (CYBCONF)* (pp. 293–298). IEEE.

Kwisthout, J., & Rooij, I. (2013). Bridging the gap between theory and practice of approximate bayesian inference. *Cognitive Systems Research, 24*, 2–8. https://doi.org/10.1016/j.cogsys.2012.12.008

Ladd, T. D., Jelezko, F., Laflamme, R., Nakamura, Y., Monroe, C., & O'Brien, J. L. (2010). Quantum computers. *Nature, 464*(7285), 45–53. https://doi.org/10.1038/nature08812

Landgrebe, J., & Smith, B. (2022). *Why machines will never rule the world: Artificial intelligence without fear*. Taylor and Francis.

Lee, N. (2019). Beauty is in the AI of the beholder: Artificial and superintelligence. *The Transhumanism Handbook,* (pp. 153–173). Springer.

Li, S. S., Long, G. L., Bai, F. S., Feng, S. L., & Zheng, H. Z. (2001). Quantum computing. *Proceedings of the National Academy of Sciences, 98*(21), 11847–11848.

Li, J. A., Dong, D., Wei, Z., Liu, Y., Pan, Y., Nori, F., & Zhang, X. (2020). Quantum reinforcement learning during human decision-making. *Nature Human Behaviour, 4*(3), 294–307. https://doi.org/10.1038/s41562-019-0804-2

Liu, Y., Arunachalam, S., & Temme, K. (2021). A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics, 17*(9), 1013–1017. https://doi.org/10.1038/s41567-021-01287-z

Lloyd, S. (2000). Ultimate physical limits to computation. *Nature, 406*(6799), 1047–1054.

Lucas, J. R. (1961). Minds, machines and gödel1. *Philosophy, 36*(137), 112–127.

Mack, C. A. (2011). Fifty years of Moore's law. *IEEE Transactions on Semiconductor Manufacturing, 24*(2), 202–207.

Markov, I. L. (2014). Limits on fundamental limits to computation. *Nature, 512*(7513), 147–154.

Martin-Delgado, M. (2012). On quantum effects in a theory of biological evolution. *Scientific Reports, 2*(1), 1–8.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine, 27*(4), 12. https://doi.org/10.1609/aimag.v27i4.1904

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics, 5*, 115–133.

McGinn, C. (1991). *The problem of consciousness: Essays toward a resolution*. Blackwell.

Mikki, S. (2023). Artificial general intelligence and noncomputability: A dynamical framework. *Journal of Artificial Intelligence and Consciousness, 10*(01), 71–101.

Miller, A. (2019). The intrinsically linked future for human and Artificial Intelligence interaction. *Journal of Big Data, 6*(1), 38.

Minsky, M. (1988). *Society of mind*. Simon and Schuster.

Moses, S. A., Baldwin, C. H., Allman, M. S., Ancona, R., Ascarrunz, L., Barnes, C., Bartolotta, J., Bjork, B., Blanchard, P., Bohn, M., Bohnet, J. G., Brown, N. C., Burdick, N. Q., Burton, W. C., Campbell, S. L., Campora, J. P., Carron, C., Chambers, J., Chan, J. W., … Pino, J. M. (2023). A race-track trapped-ion quantum processor. *Physical Review X*, *13*(4), 041052. https://doi.org/10.1103/PhysRevX.13.041052

Möller, M., & Vuik, C. (2017). On the impact of quantum computing technology on future developments in high-performance scientific computing. *Ethics and Information Technology, 19*, 253–269.

Moore, G. (1965). Cramming more components onto integrated circuits. *Electronics Magazine, 38*(8), 114.

Moore, S. (2020). Rapid scale-up of commercial ion-trap quantum computers. *IEEE Spectrum*. Retrieved May 18, 2024, from https://spectrum.ieee.org/commercial-iontrap-quantum-computers-showing-rapid-scaleup

More, M. (2013). The Philosophy of Transhumanism. In *The transhumanist reader* (pp. 3–17). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118555927.ch1

Moret-Bonillo, V. (2015). Can artificial intelligence benefit from quantum computing? *Progress in Artificial Intelligence, 3*, 89–105.

Morrison, R. (2023). Quantinuum H-Series quantum computer accelerates through 3 more performance records for quantum volume. *quantinuum*. Retrieved May 18, 2024, from https://www.quantinuum.com/news/quantinuum-h-series-quantum-computer-accelerates-through-3-more-performance-records-for-quantum-volume-217-218-and-219

Mulgan, G. (2018). *Big mind: How collective intelligence can change our world*. Princeton University Press.

Müller, V. C. (2011). On the possibilities of hypercomputing supertasks. *Minds and Machines, 21*, 83–96. https://doi.org/10.1007/s11023-011-9222-6

Müller, V. C., & Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. In V. C. Müller (Ed.), *Fundamental issues of artificial intelligence* (pp. 555–572). Springer International Publishing. https://doi.org/10.1007/978-3-319-26485-1_33

Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1690), 20150192.

Németi, I., & Dávid, G. (2006). Relativistic computers and the Turing barrier. *Applied Mathematics and Computation, 178*(1), 118–142. https://doi.org/10.1016/j.amc.2005.09.075

Nian, R., Liu, J., & Huang, B. (2020). A review on reinforcement learning: Introduction and applications in industrial process control. *Computers & Chemical Engineering, 139*, 106886. https://doi.org/10.1016/j.compchemeng.2020.106886

Nielsen, M. A. (2002). A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A, 303*(4), 249–252.

Nielsen, M. A., & Chuang, I. L. (2010). *Quantum computation and quantum information: 10th anniversary edition*. Cambridge University Press. https://doi.org/10.1017/CBO9780511976667

Ord, T., & Kieu, T. D. (2005). The diagonal method and hypercomputation. *British Journal for the Philosophy of Science, 56*(1), 147–156. https://doi.org/10.1093/phisci/axi108

Penrose, R. (1989). *The Emperor's new mind: Concerning computers, minds, and the laws of Physics*. Oxford University Press.

Penrose, R. (1994). *Shadows of the mind* (Vol. 4). Oxford University Press.

Picard, R. W. (2000). *Affective computing*. MIT Press.

Piccinini, G. (2011). The Physical Church-Turing Thesis: Modest or Bold? *The British Journal for the Philosophy of Science, 62*(4), 733–769.

Pintner, R. (1923). *Intelligence testing: Methods and results*. H. Holt.

Pitowsky, I. (1990). The physical church thesis and physical computational complexity. *Iyyun: The Jerusalem Philosophical Quarterly, 39*, 81–99.

Preskill, J. (2012). *Quantum computing and the entanglement frontier (arXiv:1203.5813)*. arXiv. https://doi.org/10.48550/arXiv.1203.5813

Putnam, H. (1967). Psychological predicates. *Art mind and Religion, 1*, 37–48.

Roitblat, H. L. (2020). *Algorithms are not enough: Creating general artificial intelligence*. MIT Press.

Rønnow, T. F., Wang, Z., Job, J., Boixo, S., Isakov, S. V., Wecker, D., ... & Troyer, M. (2014). Defining and detecting quantum speedup. *Science, 345*(6195), 420–424.

Russell, S. J., Norvig, P., Canny, J. F., Malik, J. M., & Edwards, D. D. (2003). *Artificial intelligence: A modern approach* (Vol. 2). Prentice hall.

Ryan-Anderson, C., Bohnet, J. G., Lee, K., Gresh, D., Hankin, A., Gaebler, J. P., Francois, D., Chernoguzov, A., Lucchetti, D., Brown, N. C., Gatterman, T. M., Halit, S. K., Gilmore, K., Gerber, J. A., Neyenhuis, B., Hayes, D., & Stutz, R. P. (2021). Realization of real-time fault-tolerant quantum error correction. *Physical Review X, 11*(4), 041058. https://doi.org/10.1103/PhysRevX.11.041058

Saggio, V., Asenbeck, B. E., Hamann, A., Strömberg, T., Schiansky, P., Dunjko, V., ..., Walther, P. (2021). Experimental quantum speed-up in reinforcement learning agents. *Nature*, *591*(7849), 229–233. https://doi.org/10.1038/s41586-021-03242-7

Saki, A. A., Alam, M., & Ghosh, S. (2019). *Study of decoherence in quantum computers: A circuit-design perspective (arXiv:1904.04323)*. arXiv. https://doi.org/10.48550/arXiv.1904.04323

Salisbury, D. (2022). A history of observables and Hamilton–Jacobi approaches to general relativity. *The European Physical Journal H, 47*(1), 7.

Sandberg, A. (1999). The physics of information processing superobjects: Daily life among the Jupiter brains. *Journal of Evolution and Technology*, *5*(1).

Sandberg, A., & Bostrom, N. (2006). Converging cognitive enhancements. *Annals of the New York Academy of Sciences, 1093*(1), 201–227.

Savitch, W. J. (1970). Relationships between nondeterministic and deterministic tape complexities. *Journal of Computer and System Sciences, 4*(2), 177–192. https://doi.org/10.1016/S0022-0000(70)80006-X

Scaruffi, P. (2018). *Intelligence is not artificial - expanded edition: A history of artificial intelligence and why the singularity is not coming any time soon*. CreateSpace Independent Publishing Platform.

Schrödinger, E. (1926). Quantisierung als eigenwertproblem. *Annalen Der Physik, 385*(13), 437–490.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*(3), 417–424.

Šekrst, K., & Skansi, S. (2021). *Is complexity important for philosophy of mind? (arXiv:2112.03877)*. arXiv. https://doi.org/10.48550/arXiv.2112.03877

Shagrir, O., & Pitowsky, I. (2003). Physical hypercomputation and the Church-Turing thesis. *Minds and Machines, 13*(1), 87–101. https://doi.org/10.1023/A:1021365222692

Shor, P. W. (1997). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing, 26*(5), 1484–1509. https://doi.org/10.1137/s0097539795293172

Sipser, M. (1996). *Introduction to the theory of computation* (1st ed.). International Thomson Publishing.

Smale, S. (1997). Complexity theory and numerical analysis. *Acta Numerica,6*, 523–551. https://doi.org/10.1017/S0962492900002774

Sriarunothai, T., Wölk, S., Giri, G. S., Friis, N., Dunjko, V., Briegel, H. J., & Wunderlich, C. (2018). Speeding-up the decision making of a learning agent using an ion trap quantum processor. *Quantum Science and Technology, 4*(1), 015014.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. CUP Archive.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Swami, Y., & Gill, A. (2011). Recognition of Quantum Computers vs. Classical Computers. Proceedings of the National Electrical Engineering Conference (NEEC-2011) on Powerand Energy Systems, Dheli.

Syropoulos, A. (2008). *Hypercomputation: Computing beyond the Church - Turing barrier*. Springer.

Tallis, R., & Aleksander, I. (2008). Computer models of the mind are invalid. *Journal of Information Technology, 23*(1), 55–62. https://doi.org/10.1057/palgrave.jit.2000128

Tegmark, M. (2007). The mathematical universe. *Foundations of Physics, 38*(2), 101–150. https://doi.org/10.1007/s10701-007-9186-9

Trueblood, J., Yearsley, J., & Pothos, E. (2017). A quantum probability framework for human probabilistic inference. *Journal of Experimental Psychology: General, 146*(9), 1307. https://doi.org/10.1037/xge0000326

Turing, A. M. (1937). On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, *s2-42*(1), 230–265. https://doi.org/10.1112/plms/s2-42.1.230

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*, 433–460.

Ulam, S., Kuhn, H. W., Tucker, A. W., & Shannon, C. E. (2013). *John Von Neumann, 1903–1957* (pp. 235–269). Harvard University Press.

van De Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods, 22*(2), 217.

Wang, H. (2016). *From mathematics to philosophy (Routledge revivals)*. Routledge.

Watrous, J. (2008). *Quantum computational complexity (arXiv:0804.3401)*. arXiv. https://doi.org/10.48550/arXiv.0804.3401

Welch, P. (2006). The extent of computation in malament-hogarth spacetimes. *The British Journal for the Philosophy of Science, 59*(4), 659–674. https://doi.org/10.1093/bjps/axn031

Weschsler, D. (1971). Concept of collective intelligence. *American Psychologist, 26*(10), 904.

Wichert, A., Moreira, C., & Bruza, P. (2020). Balanced quantum-like bayesian networks. *Entropy, 22*(2), 170. https://doi.org/10.3390/e22020170

Wiedermann, J. (2012). A computability argument against superintelligence. *Cognitive Computation, 4*(3), 236–245. https://doi.org/10.1007/s12559-012-9124-9

Wolpert, D. H., & Macready, W. G. (1995). *No free lunch theorems for search* (Vol. 10). Technical Report SFI-TR-95-02-010, Santa Fe Institute.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82.

Yampolskiy, R. V. (2018). The singularity may be near. *Information, 9*(8), 190.

Yang, J. M. (2022). Law of Total Probability in Quantum Theory and its application in Wigner's friend scenario. *Entropy, 24*(7), 903.

Zheng, P., & Akhmad, M.-A. (2017). How change agencies can affect our path towards a singularity. In V. Callaghan, J. Miller, R. Yampolskiy, & S. Armstrong (Eds.), *The technological singularity: Managing the journey* (pp. 87–101). Springer. https://doi.org/10.1007/978-3-662-54033-6_4

Ziegler, M. (2005). Computational power of infinite quantum parallelism. *International Journal of Theoretical Physics, 44*(11), 2059–2071.