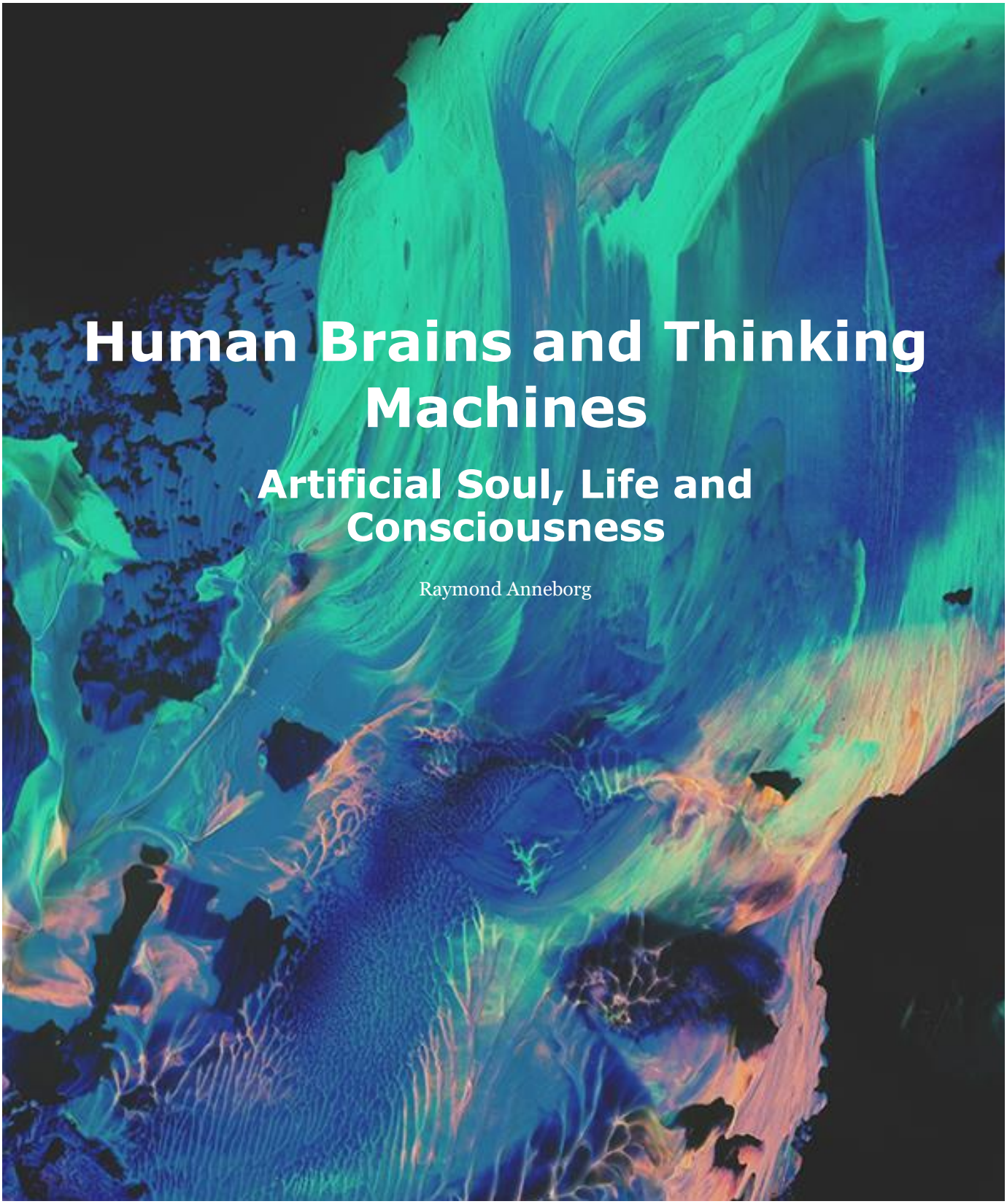
UMEÅ UNIVERSITY



Human Brains and Thinking Machines

Artificial Soul, Life and Consciousness

Raymond Anneborg

Bachelor thesis, 15 hp

Philosophy C, 30 hp

Spring term 2022

Table of Contents

Human Brains and Thinking Machines

Artificial Soul, Life and Consciousness

Raymond Anneborg

Umeå University

Abstract: In this paper, I examine if strong artificial intelligence can be achieved or not. Can machines have a mind, be conscious, think and have subjective experiences, just like a human? I analyze David Chalmers arguments supporting the possibility of strong AI and conclude that his emulation argument and principle of organizational invariance is not a sufficient condition for strong AI. Instead, I defend the thesis that *life* is a necessary condition for any conscious agent, human or machine (or other), to have a mind, be able to think and have subjective experiences. I revisit the ideas of the soul and of vitalism and the need for a life force energy, an élan vital as introduced by Henri Bergson. In the investigation of life I also examine if strong artificial life can be achieved or not, since this would be a prerequisite for strong AI.

Keywords: Strong Artificial Intelligence, Strong Artificial Life, Vitalism, Soul, Life, Consciousness, Mind, Human Brain, Thinking Machine

"My brain is only a receiver, in the Universe there is a core from which we obtain knowledge, strength and inspiration. I have not penetrated into the secrets of this core, but I know that it exists." – Nikola Tesla

"The day science begins to study non-physical phenomena, it will make more progress in one decade than in all the previous centuries of its existence." – Nikola Tesla

1. Introduction

In this paper, I will investigate *if strong artificial intelligence (AI) can be achieved or not*. The purpose is to show that strong AI cannot be achieved, because the AI systems that we develop will be non-living inanimate objects, and I will argue that ***life*** is a necessary condition for any system to have a mind, consciousness and subjective experiences. I will link this discussion to the possibility of strong artificial life (A-Life) as this would be a possible path to strong AI. My focus will be on what is *artificially possible*, that is what can be *created by man* rather than by nature or what might be metaphysically possible.

Artificial intelligence (AI) is an interdisciplinary subject spanning many different disciplines like computer science, philosophy, psychology, neuroscience, mathematics, economics, linguistics, control theory, engineering and cybernetics. There is no agreed upon definition of what we mean by AI. However, the following definition is suggested in the main textbook for AI ("Artificial Intelligence – A Modern Approach" (Russell & Norvig, 2020)):

> "The field of artificial intelligence, or AI, is concerned with not just understanding but also building intelligent entities – machines that can compute how to act effectively and safely in a wide variety of novel situations." (Russell & Norvig, 2020, p. 1).

AI is usually described with the different areas of application where it is used. This comprise of mainly different forms of machine learning (ML), natural language processing (NLP), computer vision and robotics.

Strong AI "seeks to create artificial persons: machines that have all the mental powers we have, including phenomenal consciousness" (Bringsjord & Naveen Sundar, 2020). To create this type of *thinking machines* was the original objective of AI when John McCarthy (1927-2011) at the 1956 Dartmouth summer workshop where the field of AI was born coined the term artificial intelligence. McCarthy was not the first to research the possibility of thinking machines though. Alan Turing (1912-1954) had already in his classical paper from 1950, "Computing machinery and intelligence" (Turing, 1950), introduced a thought experiment, the imitation game, which we today call the Turing test, in order to find a way to determine whether a machine can think or not.

I will also in this paper analyze David Chalmers arguments provided for the prospect of thinking machines. Chalmers has argued that strong AI and thinking machines are possible. I will review his arguments and argue that they do not provide sufficient support for strong AI. Instead, I will revisit *vitalism* and propose the thesis that *life* is a necessary prerequisite for any conscious thinking agent. I will argue that an AI machine (i.e. an "intelligent" robot) is an inanimate (programmed) object that lack life, which then lead to that a machine, cannot think or be a conscious agent, and thus thinking machines and strong AI cannot be achieved.

Vitalism suggest that, "living organisms are fundamentally different from non-living entities because they contain some non-physical element or are governed by different principles than

are inanimate things" (Bechtel & Richardson, 1998). This non-physical (or immaterial) life giving element is sometimes also referred to as *the spirit* (Latin: spiritus) of a living organism. That there is some kind of life giving force or *élan vital* (the spark of life) as suggested by the French philosopher Henri Bergson (1874–1948) (Bergson, 1911) is not a new idea. A vital force is something, which is present in many (eastern) religions where it have names like Qi, Chi and Prana that represents a *life force energy*. In discussing this élan vital, I will also touch upon the related topic of *the soul* (Latin: anima) and investigate their relationship, including the relationship to mind and consciousness. Note that the purpose of this paper is not to prove that vitalism is true. I will use it as a tool to argue that humans are more than machines.

Life is the subject for the study of biology, but as with AI there is no agreed upon definition of what life is. Like AI it is what is called a *cluster concept* of a set of properties that we associate with life, like self-reproduction, evolution, and metabolism and so on. Life can also be viewed both at an individual level and at a population level. Properties like self-reproduction is for example necessary at the population level for life, but not necessary at the individual level, where we can have individuals that are sterile and cannot reproduce, but are still alive. In this paper I will focus on *life associated with individuals* since when we talk about minds, consciousness and intelligence we are primarily associating these properties to an individual, even if we can also talk about collective intelligence and so on.

Artificial Life or **A-Life**[1] for short is the study of artificial systems that exhibit the behavioral characteristics of natural living systems, through the ***synthesis*** or ***simulation*** using computational (software), robotic (hardware), and/or physicochemical (wetware) means. The founder of this movement Christopher Langton defines A-Life as *life made by man rather than by nature* (Langton, 1986). I will review in this paper also if A-Life can be achieved and then with the focus on the claims for **strong A-Life** which suggests that our technological capabilities have brought us to the point where we are on the verge of actually creating living artifacts, based on the assumption that biological processes can be viewed as computational processes which can be implemented in a computer or artificial system. This would according to my thesis be a prerequisite for strong AI and possible path to achieve strong AI. I will argue that strong A-Life cannot be achieved and thus neither can strong AI.

[1] More information can be found at https://alife.org/.

The outline of this paper is that I will first provide some more background on strong AI, and then outline the arguments for strong AI provided by Chalmers. I will then provide some more background on vitalism and discuss how the concepts of soul, life and consciousness are related and how they have been developed throughout history. I will also provide more background on artificial life and what it aims to achieve. After this, I will argue against Chalmers arguments for AI and suggest that it is not providing sufficient support for his conclusions. I will then defend my thesis that life is a necessary condition for any conscious agent, human or machine (or other), to be able to think and have subjective experiences. I will wrap up this paper with a summary of my conclusions from this investigation.

2. Background

In this section, I will first provide some more background on strong AI, and then outline the arguments for strong AI provided by Chalmers. I will then provide some more background on vitalism and discuss how the concepts of soul, life and consciousness are related and how they have been developed throughout history. I will also provide more background on artificial life and what it aims to achieve.

2.1 Strong Artificial Intelligence

The discussions on machine intelligence (and risks related to it) have been ongoing for a long time. Alan Turing was among the first to discuss these questions in his classic paper "Computing machinery and intelligence" (Turing, 1950) from where we have the Turing test, which still no one has managed to succeed with, even though we are getting closer to solve the Turing test today. However, Turing suggested that it would be problematic to try to define what we mean by a "machine" and "think".

> "I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd." (Turing, 1950, p. 433)

This is why he proposed to replace the question “Can machines think” with his “imitation game” (the Turing test), which suggest that if a person by questioning two parties, one human and one machine/computer, cannot conclude who is who, then we can say that we have an intelligent thinking machine.

John Searle, with his famous thought-experiment the Chinese room, argues that “[…] only a machine could think, and only very special kinds of machines, namely brains and machines with internal casual powers equivalent to those of brains” (Searle, 1980, p. 417). Searle also argues that “instantiating a computer program is never by itself a sufficient condition for intentionality” (Searle, 1980, p. 417) and that a machine is “observer-dependent” (Searle, 1990, pp. 26-28) and have no understanding at all of the symbols it manipulates. He then says that “the explanation of how the brain produces intentionality cannot be that it does it by instantiating a computer program” and then concludes that “any attempt literally to create intentionality artificially (strong AI) could not succeed just by designing programs but would have to duplicate the causal powers of the human brain” (Searle, 1980, p. 417). As we will see this reasoning is in line with the arguments from David Chalmers, even though they come to opposite conclusions regarding the possibility of strong AI.

Chalmers himself devotes a whole chapter for the topic of strong AI in his book “The Conscious Mind” (Chalmers, 1996, pp. 313-332). There he asks: “Could a machine be conscious? Could an appropriately programmed computer truly possess a mind?” (Chalmers, 1996, p. 313). The next section will provide an overview of his argument for strong AI.

2.2 Chalmers argument for Strong AI

Chalmers defense of strong AI rests on his principle of **organizational invariance**, which says that for any system with conscious experiences, a system with the same fine-grained functional organization will have qualitatively identical conscious experiences. He supports this claim with a thought-experiment where neuron by neuron in a human brain is replaced by a functionally identical silicon chip until the whole brain have been replaced and we instead have an emulated human brain in the form of a machine brain. This would then constitute an implementation of strong AI. I will argue that the principle of organizational invariance is not sufficient for conscious experience.

The possibility to emulate a human brain[2] has been formalized by Chalmers in his paper on the Singularity[3] (Chalmers, 2010). Here it is called the emulation argument, and it is stated as follows:

> "(i) The human brain is a machine.
>
> (ii) We will have the capacity to emulate this machine (before long).
>
> (iii) If we emulate this machine, there will be AI.
>
> ———-
>
> (iv) Absent defeaters, there will be AI (before long)." (Chalmers, 2010, p. 14)

Here AI mean Artificial General Intelligence (AGI)[4] or human-level intelligence. Chalmers argue that "Every organ of the body appears to be a machine: that is, a complex system comprised of law-governed parts interacting in a law-governed way. The brain is no exception." (Chalmers, 2010, p. 14). He suggests that we can conclude this from what we know of biology and physics (without any specific reference though). This tie into his previous principle of organizational invariance, which suggests that, any system that is emulated with the same fine-grained functional organization will have the same causal powers leading to consciousness and subjective experiences. If this system is a human brain that is being emulated into a combinatorial-state automata (CSA)[5], based on AI algorithms then the resulting system will like a human brain also have subjective conscious experiences, thus we will have strong AI and thinking machines.

Chalmers states that "What is most relevant to the explanation of the behavior of a complex cognitive system is the abstract causal organization of the system, and computational formalisms provide an ideal framework within which this sort of organization can be described and analyzed" (Chalmers, 1996, pp. 320-321). From this statement together with the principle

---

[2] For those who want to dig into the details about brain emulations there is a more detailed report by Sandberg and Bostrom that elaborates on the topic of Whole Brain Emulation (WBE) (Sandberg & Bostrom, 2008). Bostrom as one possible path to superintelligence also discusses this topic (Bostrom, 2014, pp. 35-43). Robin Hanson has also explored the topic of brain emulation by robots in his book "The Age of Em" (Hanson, 2016).

[3] The technological Singularity is the idea that AI will reach human-level intelligence and then be able to recursively self-improve on its own to become superintelligent and beyond the control of humans (Müller, 2021).

[4] AGI is a topic which has been challenged. There are people that argue that there is no such thing as AGI. They say that humans have multiple intelligences and not one general intelligence. See (Hawkins, 2021) and (Gardner, 1999) for more details.

[5] A Turing Machine is one example of a CSA. See (Chalmers, 1996, pp. 315-322) for a more detailed explanation of what a CSA is.

of organizational invariance, Chalmers put forward his thesis of strong artificial intelligence that says:

> "It follows that for a given conscious system M, its fine-grained functional organization can be abstracted into a CSA M, such that any system that implements M will realize the same functional organization, and will therefore have conscious experiences qualitatively indistinguishable from those of the original system" (Chalmers, 1996, p. 321).

Chalmers concludes that this will be dependent on whether or not cognitive dynamics are computable or not. At the same time, in his analysis of the Singularity, he says that mentality and consciousness are immaterial non-physical processes, and "If these processes cannot be emulated or artificially created, then it may be that human-level AI is impossible" (Chalmers, 2010, p. 16). I will argue that Chalmers discussion and his arguments do not provide sufficient support for strong AI.

2.3 Vitalism – Soul, Life and Consciousness

Vitalism holds that biological phenomena can be viewed as vitalistic, that a vital spirit is a substance that gives life to bodies. It is also a position that has been developed in response to the mechanistic view – "Mechanistic explanations of natural phenomena were extended to biological systems by Descartes and his successors. Descartes maintained that animals, and the human body, are 'automata', mechanical devices differing from artificial devices only in their degree of complexity"[6] (Bechtel & Richardson, 1998).

However, the discussion on life goes all the way back to Aristotle. In De Anima (On the Soul) (Aristotle, 2018) he suggested that life is a nested hierarchy of capacities, such as metabolism, sensation, and motion. It is interesting to note here the usage of the word anima (or Greek psyche) in the context of describing what life is. Anima actually covers not just the soul, but also life, consciousness and being. Soul is identified as a substantial entity believed to be that in each person, which lives, feels, thinks and wills, or the spiritual and emotional part of a person that animate existence or life.

[6] We should note here though that Descartes also suggested that there is an immaterial "thinking substance" (res cogitans) in his substance dualism.

The concept of consciousness was introduced quite late by John Locke in his “An Essay Concerning Human Understanding” (Locke, 1997) when the talking about the soul became out of fashion he replaced it with consciousness instead. Locke defines the self as "that conscious thinking thing, (whatever substance, made up of whether spiritual, or material, simple, or compounded, it matters not) which is sensible, or conscious of pleasure and pain, capable of happiness or misery, and so is concerned for itself, as far as that consciousness extends" (Locke, 1997).

If we look at Chalmers quote, which said that we should view every organ of the body as a machine, including the human brain, it seems that the mechanistic view has regained its position today. The key challenge with this position is to explain how a complex living organism can result from just matter and motion. A new life starts with cells that are divided, forming a body (the phenotype) according to the genetic DNA blueprint (the genotype or genome), but how can this become a living conscious organism?

Vitalism has fallen out of favor. It is viewed as a position that proposes an unknowable factor in explaining life, which cannot be falsified. Also today’s materialism, which is advocated by science and spearheaded by neuroscientists and biologists, when it comes to body, mind and consciousness, think that modern theories related to Darwinian evolution and genetics can explain life and thus have “demystified” life. I do not agree with this claim and will argue against it in this paper.

2.4 Artificial Life (A-Life)

The field of Artificial Life (A-Life) is devoted to studying the scientific, technological, artistic, philosophical, and social implications of creating "living" artifacts. One of the main objectives of A-Life is to seek to increase the role of synthesis in the study of biological phenomena. However, there are also proponents for **strong A-Life**, like Elliott Sober, which do not limit these studies to just synthesis or simulation, but who think that life can actually be created by man. I will argue against this position as well in this paper.

Most proponents of A-Life are also functionalists, which in this context means that their view is that the properties of life can be described as *informational concepts*, which can be computed, or that biological processes can be viewed as computational processes. They see things like autonomy, adaptation, evolution and so on as structure, process and control that can be informationally expressed. Based on this they also draw the conclusion that “If living

self-organization is definable in logical terms, then a virtual 'creature' implemented in computer memory that satisfied these abstract criteria – whatever they are – would be genuinely alive." (Boden, 2006). This would then be support for the claim of strong A-Life. However, we also have *metabolism* as a key property of life, and metabolism concerns *energy* and not information. Functionalist A-Life proponents try to avoid this topic, but when they do address it they suggests that a computer also consumes physical energy and that metabolism can actually also be functionally defined. I will argue, when examining which life properties that apply to machines, that metabolism is something that is only related to living systems like humans and that machines do not have metabolism.

3. Investigation

In this section, I will present and defend my thesis, where I suggest that ***life*** is a necessary condition for any conscious agent, human or machine (or other), to be able to think and have subjective experiences. I will argue against proponents of strong AI, like David Chalmers, and also against proponents of strong artificial life, like Christopher Langton and Elliott Sober. I will argue that life is something that can only be attributed to living biological organic systems, like humans and animals.

3.1 The Life Thesis of Mind

In this section I will outline my thesis and argue that ***life*** is a *necessary* condition, but not a *sufficient* condition, for a system to have a mind, be conscious and have subjective experiences. However, what do we mean by life?

Biology is the science of life, but biologists have no precise definition of what life is, they cannot provide us with a description of the nature of life. We saw from Aristotle that life could be seen as a hierarchy of capacities, like metabolism, hearth rhythm, blood pressure and other vital biological processes. Other properties often related to life are self-organization, emergence, autonomy, growth, development, reproduction, evolution, and responsiveness. There are probably more properties that can be attributed to life as well. The question is then if an AI can emulate these properties or implement biological processes computationally, that is to say can there be strong artificial life? I will argue that strong artificial life cannot be achieved, and thus neither can AI consciousness or strong AI be achieved. I will also suggest

that in order for a mind to be created we need a self-aware sentient (complex and organic, animated) system. A single cell can be considered a living organic system[7], indeed the very fundamental building block for life. A heart can be viewed like this as well. It can be kept alive. It can be transplanted from one person to another and so on. However, neither a cell nor a heart would be considered to have a mind. Such entities will not have a will or desires, imagination or thought. A cell or heart in itself will not experience pain or happiness, i.e. they will not be sentient. It is first when we have a complex system like an animal including humans that we say that these entities have a mind. So whether or not we attribute a mind to a living system is a matter of the complexity of self-organized matter.

To formalize my arguments above I will propose a thesis that suggest that life is a necessary property for a system to be able to have a mind and that a mind is needed to have conscious subjective experiences, which by definition is needed for there to be strong AI. This was once seen as an obvious relationship, but since "mid-century the link between life and mind – though still widely accepted, even taken for granted – wasn't explicitly considered" (Boden, 2006, p. 1430). My aim is to reinstate this link.

The Life thesis of Mind

Only a living animated system can have a mind.

I will support this thesis by arguing that *all* the minds that we know about are found in living system/organisms, like humans and animals. At the same time *no* minds that we know about can be found in any non-living system/machine. Since AI can be viewed as a type of machine, i.e. a programmed computer or robot, this then leads to that we will not find any mind connected to AI. Therefore there will be no AI mental activities, no thinking, no sensations or subjective experiences and thus **strong AI cannot be achieved**. Below I have formalized this argument.

P1: All minds are found in living systems.

P2: Non-living systems do not have minds.

P3: Machines, including AI, are non-living systems.

[7] Cellular automata is one of the approaches used in A-Life to study natural life. See (Langton, 1986).

———

C1: Machines do not have minds.

C2: Strong AI cannot be achieved.

In the remainder of this paper I will argue for each of my premises and conclusions, and investigate potential objections from proponents of strong AI and strong artificial life. However, before I do that I will provide you with a thought-experiment to show that **humans are more than machines**, that we are not just a law-governed mechanical body. The thought-experiment will be used in the following discussions supporting and defending my thesis and argument.

3.2 The RoboEinstein Thought-Experiment

In this section I will provide a thought-experiment in order to show that humans are more than just mechanical automata or machines made up of just a complex material body with parts in the form of law-governed machine organs, as suggested by Chalmers (and Descartes) in my description of Chalmers argument for strong AI above. I will suggest that ***life*** is the distinguishing property between man and machine. So, let us imagine the following:

- We have access to Einstein's brain.
- We have the ability to clone a brain into an identical copy.
- We also have the ability to do whole brain emulations (WBE) and create a functionally identical isomorph of a brain (based on Chalmers principle of organizational invariance).
- We have the technology and engineering skills as well as medical skills needed to be able to do a whole brain transplant.

Now, given the above, we could create an identical copy of Einstein's brain. Let us put that in a vat for now. We could also create a silicon isomorph of Einstein's brain (which would be functionally identical to Einstein's biological brain that we just copied). Actually, let us make two such silicon isomorph copies of Einstein's brain, and let us put one of them into RoboEinstein, a machine and our strong AI project, and let's store the other copy for now.

Let us further imagine that Einstein and RoboEinstein meet and that they end up having an argument over whether or not quantum physics will destroy the theory of general relativity. This lead to a duel between them where unfortunately both of them are killed by a bullet through their heads. So this is bad news. Both are now dead. Luckily we have access to the copies we made of their brains, both Einstein's biological brain that we stored in the vat and RoboEinstein's second silicon isomorph copy of Einstein's brain. A team of brain surgeons now get started with Einstein and carefully replaces the destroyed brain with the copy of Einstein's brain that we had in the vat and connects it to Einstein's body so that after the operation we have again Einstein in the same physical state as before the duel. A team of robot engineers then do the same thing with RoboEinstein and replaces the destroyed silicon isomorph brain with the second copy we made, resulting in that also RoboEinstein after this repair now is in the same physical state as before the duel. The engineers then *recharge* RoboEinstein and *turns him on*. He wakes up wondering why he does not have any memories of the last 24 hours since the duel. However, when the team of surgeons try to *revive* Einstein, they are unsuccessful and Einstein remains dead.

Now, what conclusions can we draw from this? Einstein and RoboEinstein are isomorphs with the same internal causal powers and the same complexity of their physical constitution. It should be noted here that the order of complexity does not really matter. The human physical constitution of self-organized matter could be a million times more complex compared to what it is, but the result would still be the same, Einstein would remain dead and RoboEinstein would be repaired and brought back to "life". The difference between the two is that Einstein is made up of a biological organic carbon cell-based substrate and RoboEinstein is made up of a metal and silicon chip-based substrate. Functionally they are the same. Whether or not they support organizational invariance can be discussed, since from a functionalist perspective multiple realizability do not require this. Then there is also the difference that Einstein, the human organic person, has metabolism which is about energy rather than information, physical constitution and functionality, while RoboEinstein do not have any metabolism.

So, how come that Einstein could not be brought back to life when RoboEinstein could be? There seem to be something more that Einstein had but RoboEinstein do not have. Einstein was an individual *living* system, while RoboEinstein is a non-living machine. Now, if a human is nothing more than a physical machine subject to natural laws, as suggested by

Chalmers, or an automata as suggested by Descartes (for the human body), then replacing the destroyed brain of Einstein with an identical copy should have succeeded in the same way as it did when replacing the brain of RoboEinstein, but it did not. Therefore we must conclude that a human is more than a machine. The *living* Einstein had what I would call a *vital life force*, an *élan vital*. You could call it something else if you do not support vitalism, however the fact remains, *life* left Einstein when he was killed, and he lost his vital life force (his soul departed). This is why he could not be *reanimated* (his soul could not be brought back) and brought back to life. This also shows that if *the difference* between Einstein and RoboEinstein is ***life*** (in whatever way you chose to explain life), and Einstein had life, then RoboEinstein cannot have life (otherwise this difference would not exist). This means that RoboEinstein is a non-living system, a machine, and thus cannot have a mind and therefore cannot have consciousness or subjective experiences either, and thus **strong AI cannot be achieved**.

Before we move on to discuss my argument supporting the Life thesis of Mind let us look at a couple of potential objections against my thought-experiment above.

Let us start by looking at the objection that I do not take into account the vital biochemical physical processes within a human body, like breathing, pulse or heartbeat and so on. At the point of death these vital processes would stop and this would be the reason for us not being able to revive Einstein. However, if the human is just a law-governed mechanical machine, just like RoboEinstein, then replacing broken "parts" (like Einstein's brain) should have "repaired" (revived) Einstein so that we should have been able to "recharge" and "turn him on" again, like we could do with RoboEinstein. Therefore referring to vital processes as the cause of failure to revive Einstein is not sufficient. You could compare it to another machine like a car. A cars "vital processes"[8] would be the powering of the engine by fuel or the lubrication of parts by oil. If you would put petrol into a diesel car the engine would break down (the vital process would stop). However, if you then replace the broken engine with a new one and refuel the car with proper fuel, in this case diesel, then there would be no problem to restart the car. It would again be "alive". So, there is more to a human than to a machine, and as I have suggested the difference is *life* which is a property of a human but not of a machine.

[8] As you can see from this example it is hard to even talk about vital processes in the context of a machine, which again indicates that humans are more than machines.

Let us now turn to the neuroscientist objection. One of the champions of this objection is the neuroscientist Anil Seth whom have said that:

> "Briefly, the vitalist notion that life could not be explained in terms of biophysical mechanisms was neither directly solved (by finding the elusive 'spark of life') nor eradicated (by discovering that life does not exist). It was dissolved when biologists stopped treating life as one big scary mystery, and instead started accounting for (i.e., explaining, predicting, and controlling) the properties of living systems (reproduction, homeostasis, and so on) in terms of physical and chemical processes. We still don't understand everything about life, but what seemed at one time beyond the reach of materialism no longer does." (Seth, 2021, p. 56)

I would argue that this approach of accounting for bits and pieces of life might give us some pieces of the puzzle, but it will not provide us with the whole picture. I would also view this bits and pieces approach as a black swan argument[9]. Just because physical and chemical processes examined by neuroscientists can explain one thing, a second thing and a third thing, does not mean that we can conclude that it can explain everything. Also, Seth states "We still don't understand everything about life" (which is true) and implies that life will however be completely mapped out and explained by science *in the future*. This is very similar to Schrödinger's view in (Schrödinger, 2012) where he talks about life as being governed by laws of nature that are *yet unknown to us*. So both Seth and Schrödinger are relying on future scientific discoveries to explain life. This is of course not possible to test or falsify, since these scientific discoveries have not been made yet. I would therefore argue that this objection is very weak since it cannot be supported by any justifications for this *belief* to show that it would be true.

A second objection would be the question "why should we think that it's impossible to repair Einstein's brain and bring him back to life"? We might not have the technology for that today, but what says we will not have that in the future. Let us grant this possibility for a while. It would be quite extraordinary for humanity to develop technology and medicine to the extent that we can "repair" even a gunshot to the head which destroys the brain. It would mean that

[9] A black swan argument is an argument based on induction. We assume that all swans are white because we have only observed white swans. However, then a black swan is observed and the argument turns out to be false (which is a possibility for all empirical observation based conclusions).

we will reach a point where we can become immortal and have an eternal lifespan (since any physical injury whatever it might be that causes "death" can always be fixed and a dead person can always be brought back to life). Let us hope that when we become so advanced we have also found some way to leave the Earth and colonize space, because the human population would then grow exponentially since there would be births but no one dying. The Earth's resources would be depleted very quickly and would no longer be able to sustain the human population with food and water or space to live.

Now, we could also modify the thought-experiment so that instead of shooting Einstein and RoboEinstein to death and destroying their physical "parts", we instead remove their energy sources. This would mean that we *turn off* RoboEinstein and that we starve Einstein to death by removing food and water. When Einstein have died, this time from starvation, we can again *turn on* RoboEinstein, that is once again provide the needed energy and he will function just as before we turned him off. If we on the other hand try to provide the dead Einstein with energy again he will not start eating and drinking, and even if we forced food and water into Einstein there would be no metabolism that would convert it to energy that could revive Einstein, he would once again remain dead.

Now, let us examine each of the premises and conclusions in my argument that supports my Life thesis of Mind.

3.3 All minds are found in living systems (P1)

To support this premise I will use the scientific approach. By observation we can conclude that there is very strong empirical support for the fact that all minds are to be found in living systems. We are almost 8 billion humans living on this planet and for each and every one of us we can say that we have a mind, and this is not counting the animals (or plants). So we have a very large data set with a consistent observation result in favor of P1.

Of course the scientific approach is subject to the possibility that in the future we might suddenly observe a living conscious machine (which would then make P1 and P3 false), so it is subject to the inference problem (the black swan scenario mentioned above) of science. However, this is the skeptic's objection and if we would like to reject this premise based on the rejection of the scientific method for knowledge acquisition, then we would also have to

reject all other scientific results that we have. I do not think that many would like to take that position.

We should also note here that this example is different from the black swan argument provided by Seth and Schrödinger above. Even though we have induction used in both cases to argue the point, the data set used in my example here is very large (many billions of data points with the same type of observation result), while the Seth/Schrödinger observations consist of a comparable very limited data set of observations consisting of a handful of scientific explanations for different life processes. To say that I can explain 100 life processes with science and therefore all life processes will be possible to explain with science is not as strong as saying that 8 billion living systems have minds and therefore all living systems have minds.

So what other objections might we have against P1? There is of course solipsism, (Latin: solus "alone" and ipse "self"), which says that the only thing we can really know when it comes to minds is that my own mind exists. As for other minds we cannot be sure. It might be that the external world, including other humans with minds, is just a construct within my mind and do not exist in reality. It should first be noted here that the solipsist position cannot be falsified. The only way to know if someone has a mind or not is to be that person (or machine) and to feel oneself thinking, which is of course not possible. Alan Turing discusses solipsism when he examines "The Argument from Consciousness" against his imitation game (Turing test) (Turing, 1950, pp. 445-447). There he concludes that:

> "It is in fact the solipsist point of view. It may be the most logical view to hold but it makes communication of ideas difficult. A is liable to believe 'A thinks but B does not' whilst B believes 'B thinks but A does not'. Instead of arguing continually over this point it is usual to have the polite convention that everyone thinks." (Turing, 1950, p. 446)

This is also in line with our intuitions about what we consider living and thinking. We believe that other humans are also alive, have a mind and can think, even though we cannot justify such beliefs based on a clear definition of what we mean by life or what it feel like to be another thinking person (or machine).

Yet another objection could be to say that "all of the minds we know about are located on Earth, but no one would think being located on Earth is a necessary condition for having a mind". This is true, and it is also one of the things that A-Life researchers wants to go beyond. A-Life want to understand *life as such* wherever it might appear, so not necessarily just on Earth, even though also A-Life researchers acknowledge that it is only life on Earth that we, at least at this point in time, have access to and can study. However, this objection misses the target. The premise here do not put any constraints on where the living system that we observe is located. It could be on Earth, the Moon, on Mars or anywhere else in the Universe, our observation would be the same. We would find a mind.

A final objection would be to say "what about animals and plants, they are alive, do they also have minds"? This would however be a mistaken objection. Even if we have a living system, like I discussed above when talking about a cell or a heart, this is not a *sufficient* condition for there to be a mind. Life is a *necessary* condition, so if we find a mind, then we can be sure that we have found it in a *living* system (at least based on current empirical evidence). The question is not without merit though and there are indications that both animals and plants actually show behavior that indicate both intelligence and sentience. However, this is not the topic for this paper.

3.4 Non-living systems do not have minds (P2)

Again we can use the scientific approach and examine all different types of non-living systems that we can think of, like your coffee cup, the refrigerator you have at home, your car, or your computer which you have programmed with the most sophisticated AI algorithms that we have today, even RoboEinstein in the thought-experiment above, to see if we can find any mind. Such investigation, which also could cover several billion systems examined, would end up with the result of zero minds found.

These are inanimate objects. Here it is also interesting to see our use of language when we express ourselves about living and non-living systems. We talk about *animate* living systems and *inanimate* non-living systems. The etymology of animate and inanimate stems from anima (a current of air, wind, air, breath, the vital principle, life, soul). So, an animate living system would have a vital principle, an élan vital, soul and be alive, while an inanimate non-living system would not have this. This is also closely related to mind (Latin: animus – the

rational soul in man, intellect, consciousness, will, intention, courage, spirit, sensibility, feeling, passion, pride, vehemence, wrath, etc., the breath, life, soul).

We could also view this premise in the context of a human. If a human have died, like Einstein in our thought-experiment above, then there will be no mind anymore. All mental activities have ceased, not just the physical processes in the body. There would be no more perception of the outer world or subjective experiences of pain or pleasure. To make this even clearer we could cremate Einstein's body to make his brain completely vanish. With the physical body destroyed there would be no place to allocate a mind. So non-living systems, whether machines or dead humans (or animals, plants etc.), do not have minds.

Now, one objection here would be that which Alan Turing brought up in his response related to consciousness of machines above. How can we actually be sure that a machine do not have a mind and cannot think? The only way to be sure would be to actually be that machine, which we cannot be. Then we are back at solipsism. Would there be any other way to avoid solipsism? There have been tests proposed for examining whether or not an AI is conscious. One such test, the AI Consciousness Test (ACT), have been proposed by Susan Schneider (Schneider, 2019, pp. 46-71). To pass this test a machine/AI would have to provide a *satisfactory* response to one or more questions. Some sample ACT questions are listed below:

1. Could you survive the permanent deletion of your program? What if you learned this would occur?
2. What is it like to be you right now?
3. You learn that you will be turned off for 300 years, beginning in an hour. Would you prefer this scenario to one in which you had been turned off in the past for the same length of time? Why or why not?

More sample questions can be found in (Schneider, 2019, p. 55). The problem with this test is that it would only be able to provide behavioral indicators of consciousness. Schneider believes this is sufficient for determining whether or not an AI would be conscious. She calls it a *zombie filter*. I would argue that an ACT test like the one proposed here would not be enough to decide whether or not a non-living system has a mind, is conscious and can think. The development in Large Language Models (LLM) like Google's LaMDA and Open AI's GPT-3

show that these *programs* are becoming quite good at answering these type of questions, but still *almost* no one (see (Tiku, 2022) for the most recent example of a Google engineer saying that an AI actually have become sentient) would say that these LLM's would be alive.

3.5 Machines, including AI, are non-living systems (P3)

To support this claim we would have to examine the properties that we associate with life, like self-reproduction, evolution and metabolism, and see if we would be justified or not in assigning them to machines/AI. Therefore let us start with such a list of life properties. In this investigation I will use the list proposed by Aristotle in De Anima (DA) books II and III (Aristotle, 2018, pp. 21-69) where he discusses the definition of the soul. Below I have extracted what I believe is a complete list of the life properties provided by Aristotle:

1. Nutrition (metabolism).
2. Growth (evolution).
3. Reproduction.
4. Perception (sensation).
5. Imagination.
6. Thought.
7. Desire (wishing).
8. Self-movement (autonomy).

Aristotle calls these "psychic powers" or "soul powers". We should note here that Aristotle attribute these powers or life properties to a soul (anima) and not to a body. Another thing to note here as well is that this list of life properties have since Aristotle reappeared in different, but very similar, forms by others who try to define what we mean by life.

Since we are dealing with a list of properties it will be fuzzy. How many of these properties need to be present for us to conclude that a system is alive[10]? I have already mentioned reproduction as one example of a property which do not have to be present for a person whom we would still consider alive. Let us take RoboEinstein from the thought-experiment above and examine how many of these life properties that we could possibly grant

[10] Aristotle thought that *one* life property was enough.

RoboEinstein, and where this would lead us in terms of deciding whether we think that RoboEinstein is alive or not.

So let us start with *nutrition and metabolism*. Humans consume food and water to sustain themselves. This is then converted to energy through metabolism. Would RoboEinstein have this or something similar that we could call metabolism? As mentioned above A-Life supporters are mainly functionalists and view everything as information that can be computed. When it comes to metabolism though this is not about *information* but about *energy*. RoboEinstein, or any other machine for that matter, would also require energy, but this would be consumed by plugging in to an energy source directly, like an electric outlet or by the use of solar panels, for loading the batteries that power the machine. In my view this is an indication that a machine would not have this life property.

What about *growth and evolution*? Well, this is one of the fears related to artificial superintelligence that an AI would go into a recursive loop of self-improvement (the so called singularity) and thus evolve into a system that is way more intelligent than any human. A system that we could no longer control (the control problem) or secure that it would share our values (the value alignment problem)[11]. This is also a focus area within A-Life to study the evolution of systems by means of artificial simulations like Langton's cellular automata (Langton, 1986). Therefore it would be possible for a machine to grow and evolve, so we would have to grant this as a possible life property for RoboEinstein.

Then we have *reproduction*. Could RoboEinstein create another RoboEinstein? Most likely RoboEinstein could be programmed to assemble another RoboEinstein, a clone. We also have computer viruses that spread by copying themselves onto new computers. In this view then also reproduction would be a life property that we would have to grant RoboEinstein.

Next is *perception and sensation*. Now we do have both computer vision and natural language processing (NLP) based on sound analysis algorithms. To some extent a machine can at least detect if it has "touched" something. Smell however is still not supported by any AI algorithm and the same goes for taste. Also, perceptions (external) are linked to sensations (internal) of that which is perceived. Here we have the hard problem of consciousness and the real question of the possibility for strong AI. Could an AI be conscious, sentient and have

[11] More information on these topics can be found in (Bostrom, 2014).

subjective experiences (sensations)? I argue through-out this paper that this cannot be achieved. Thus I would not grant RoboEinstein this life property.

Imagination. Aristotle distinguished imagination from both perception and thought "for in both cases the soul discriminates and is aware of something which exists" (Aristotle, 2018, p. 52). Imagination on the other hand does not have to be related to *something which exists*, we could imagine a unicorn for example. Would imagination be a life property that RoboEinstein would have? I would say no. Even an advanced AI like RoboEinstein would still be based on a program which would set the boundaries for what such a machine could "imagine".

Now we have *thought*. Again one of the key questions. Can we have thinking machines? Well, we can have information processing based on machine perceptions. Would this constitute thinking? This is a "who came first question", the chicken or the egg. In order to determine if a system could think we would, at least according to me, have to decide if that system would be alive, but in order to decide if the system is alive we would have to answer if it can think (together with possessing other life properties). As for RoboEinstein we have already concluded that perception and sensation is not a property that we could grant RoboEinstein, and since thinking would need this property, I would suggest that we cannot grant RoboEinstein thinking either.

Can machines have *desires and wishes*? Well, most AI algorithms (i.e. reinforcement learning) are based on some *objective* that should be maximized. Would this be the same as having a desire or wish? I would say no. RoboEinstein would not desire RoboMarilyn or wish for a cold beer on a hot day.

Finally *self-movement and autonomy*. Robots can move on their own and we have today lots of debates on autonomous lethal weapons (where the decision to kill or not to kill is given to the weapon system itself). An AI system will based on the "perceptions" (information input) that it receives compute a response in terms of what to do next in order to best achieve its objective. This could to some degree be called self-movement and autonomy, so I would grant this life property to RoboEinstein.

Let us summaries our investigation into which life properties that could be granted for a machine/AI below and compare it to humans.

Life Property	Human	Machine/AI
Nutrition (metabolism)	Yes	No
Growth (evolution)	Yes	Yes
Reproduction	*Yes*	*Yes*
Perception (sensation)	Yes	*No*
Imagination	Yes	No
Thought	Yes	No
Desire (wishing)	Yes	No
Self-movement (autonomy)	Yes	Yes

Table 1 - Summary of Life Properties Analysis

As we can see machines or AI systems are, according to our analysis, lacking a lot of the life properties proposed by Aristotle. It should be noted that reproduction might not be present in all individuals and that machines/AI have some perception capabilities. Based on this analysis I would conclude that machines, including AI, are non-living systems.

So, what objections could be raised against the conclusion of this analysis? As I mentioned above since life is a cluster concept it will be a gray zone when it comes to systems that have some but not all of the life properties above. Aristotle even suggested that *one* supported life property was enough to say that a system is alive. We do have three "yes" for machines/AI so why is not this enough to say that a machine is alive? Let us look at the three life properties that we granted RoboEinstein and take some other examples where these properties would be present in a machine, but where we would definitely say that the machine/AI is not alive.

Let us first take *reproduction* and look at a computer virus, which is a program in the same way as an AI system is a program. A computer virus can reproduce itself, but we would not say that a computer virus is alive even though it have this property. It is just a program, a piece of computer code. We could then also associate this example with *evolution and growth*. As discussed in our analysis it is possible for an advanced AI system to improve or evolve its own program into a better more advanced program, but it is still just a program and we would not consider it to be alive. For *self-movement and autonomy* let us look at autonomous vehicles that is self-driving cars, trucks, construction equipment etc. These vehicles are using AI to move around in traffic on their own, so both self-movement and autonomy. However, we would not say that such a vehicle would be alive. Thus it does not seem to be the case that one, as suggested by Aristotle, or even a few life properties would be a sufficient condition for saying that a system is alive.

Then we have of course the objection from proponents of strong artificial life whom would claim that man can indeed create life in an artificial system or machine. The idea here is that "[…] computers are instances of biological processes. Here the computer is said to be alive, or to exemplify various properties that we think of as characteristic of life" (Sober, 1992, p. 750). Sober then continues to also discuss strong AI and suggests that "[…] computers are minds. This latter idea, roughly, is what usually goes by the name of strong AI" (Sober, 1992, p. 750). He then continues to argue that "human cognition is based on the idea that cognition involves computational manipulation of representations" which can be implemented by AI, and continues "Can an independent case be made for the idea that biological processes in naturally occurring organisms involve the formation and manipulation of representations? If this point can be defended for naturally occurring organisms, then to the degree that computers can form and manipulate representations in the right ways, to that degree will the AL research program appear plausible" (Sober, 1992, p. 763). So, Sober suggests that if biological processes can be view as computational processes, then strong A-Life can be achieved. He is relying on the functionalist thesis of multiple realizability (which would be the opposite of Chalmers organizational invariance). However, he also brings up problems with his thesis for strong A-Life, like digestion (nutrition and metabolism in my analysis above, and also in the RoboEinstein thought-experiment) which is a life property that does not involve the formation and manipulation of representations, since it is about extracting energy from in this case food particles. He also avoids the question about *the **nature** of life*, which he leaves to philosophers to answer (since biologist do not have any answer to this question). Instead he suggests that biologists know a lot about various specific living systems and that this is enough. In the same way he then suggests that "If a machine can extract energy from its environment, grow, repair damage to its body, and reproduce, what remains of the issue of whether it is 'really' alive? (Sober, 1992, p. 766). I have argued that the nature of life actually do matter, at least in terms of determining if a system have a mind or not, and whether a machine is "really" alive also matters. So, from my analysis of potential machine life properties and the analysis of the ones we granted a machine to have, we can conclude that this is not enough for us to attribute life to a machine. Therefore, based on these arguments, **strong A-Life cannot be achieved**.

3.6 Machines do not have minds (C1)

This conclusion follows logically from P2 and P3, so since I above have argued that P2 and P3 are true, then it also follows that C1 is true.

3.7 Strong AI cannot be achieved (C2)

Finally this is the conclusion that was the purpose of this paper to argue that **strong AI cannot be achieved**. From P1, P2 and P3 we have concluded C1. As life is needed for a system to have a mind, and a mind is needed for any mental activities including subjective experiences, and machines/AI are non-living systems, C2 also follows. Strong AI would require subjective experiences which we cannot have without a mind and without life.

Now David Chalmers has of course suggested that *organizational invariance* in a silicon isomorph, like RoboEinstein, is necessary for subjective experiences to emerge, so he would object to this conclusion. According to Chalmers for a silicon isomorph "organizational invariance is naturally necessary but not logically necessary"[12]. That is to say that if we have organizational invariance (based on a system, like a human, that have subjective experiences) then subjective experience must follow, even though this does not imply that it is metaphysically necessary that subjective experience must follow (i.e. it is possible to logically conceive of a silicon isomorph without consciousness that would not have subjective experiences – compare with Chalmers zombie argument).

In my thought-experiment above I have suggested that a human being is more than a machine (the mechanistic view proposed by Chalmers) and that the thing which distinguished a human from a machine/AI is *life*. I have shown that even if we have organizational invariance in a system, like Einstein before and after being shot and having the brain transplant, this is not sufficient for subjective experience. Einstein had subjective experiences when he was alive, but after being shot when he died to restore the organizational invariance in Einstein[13] by means of the brain transplant was not enough for Einstein to continue to have subjective experiences (because he remained dead rather than being alive again). Thus Chalmers

[12] From private email correspondence with Chalmers.

[13] This would be a *clone* rather than an *isomorph* since it would be the same substrate and not a different substrate as with RoboEinstein whom would be made of silicon rather than organic carbon based matter. However, the organizational invariance would still be there.

principle of organizational invariance is not sufficient for consciousness and subjective experiences.

It should also be noted here that organizational invariance as a necessary property for subjective experiences is in opposition to the functionalist thesis of multiple realizability which want to abstract away from the physical and suggest that any physical organizational *variance* that is functionally equivalent will realize a specific property, like properties for being alive or having a mind, consciousness and subjective experiences.

4. Conclusion

In this paper, I have investigated *if strong artificial intelligence (AI) can be achieved or not*. I have argued that strong AI cannot be achieved, because the AI systems that we develop will be non-living inanimate objects, and I have in my thesis suggested that ***life*** is a necessary condition for any system to have a mind, consciousness and subjective experiences. This discussion have been linked to the possibility of strong artificial life (A-Life) as this would be a possible path to strong AI. My focus have been on what is *artificially possible*, that is what can be *created by man* rather than by nature or what might be metaphysically possible.

I have also discussed the arguments for strong AI provided by David Chalmers, based on his emulation argument and principle of organizational invariance, and argued that they do not provide sufficient support for strong AI. I have looked that the claims for strong A-Life suggested by Elliott Sober, based on functionalism and the possibility for seeing biological processes as computational processes, and concluded that his arguments do not hold for processes like digestion/metabolism and that he avoids the question of the nature of life.

Instead I have outlined my thesis and argued that ***life*** is a *necessary* condition, but not a *sufficient* condition, for a system to have a mind, be conscious and have subjective experiences. I supported my thesis with a thought-experiment where I suggested that a human is more than a machine, and that the difference between humans and machines is life, which humans have and machines do not have. I also supported my thesis by arguing that all minds that we know of are found in *living* systems and that no minds can be found in non-living systems, like machines and AI.

I have defended my thesis against a number of objections, showing that they either force the objector into a position of skepticism where we would have to suspend judgements and conclude that we do not know anything, or a solipsist position where the only thing we can be

sure of is our own subjective experiences, which will not lead to very fruitful discussions of any kind. I also showed that objections that are referring to discoveries that might happen sometime in the future (or not) are very weak claims to refute my thesis. I have also at some length made an analysis of which life properties that we could potentially attribute to a machine and concluded that these are not enough to say that a machine would be alive, and thus **strong A-Life cannot be achieved**.

The conclusion then is that since a machine, including AI, is an inanimate non-living object, machines cannot have a mind, be conscious and have subjective experiences, and thus **strong AI cannot be achieved**.

Bibliography

Aristotle. (2018). *De Anima (On the Soul).* Oxford: Oxford University Press.

Bechtel, W., & Richardson, R. C. (1998). Vitalism. In *Routledge Encyclopedia of Philosophy.* London: Routledge.

Bergson, H. (1911). *Creative Evolution.* New York: Henry Holt and Company.

Boden, M. A. (2006). What's Life Got To Do With It? In M. A. Boden, *Mind as Machine: A history of Cognitive Science (Vols 1-2)* (pp. 1430-1443). Oxford: Oxford University Press.

Bostrom, N. (2014). *Superintelligence - Paths, Dangers, Strategies.* Oxford: Oxford University Press.

Bringsjord, S., & Naveen Sundar, G. (2020). *Artificial Intelligence (Philosophy of Artificial Intelligence).* (E. N. Zalta, Ed.) Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/artificial-intelligence/#PhilArtiInte

Chalmers, D. (1996). *The Conscious Mind - In Search of a Fundamental Theory.* Oxford: Oxford University Press.

Chalmers, D. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies (17)*, 7-65.

Gardner, H. (1999). *Intelligence Reframed - Multiple Intelligences for the 21st Century.* New York: Basic Books.

Hanson, R. (2016). *The Age of Em - Work, Love and Life when Robots Rule the Earth.* Oxford: Oxford University Press.

Hawkins, J. (2021). *A Thousand Brains: A New Theory of Intelligence.* New York: Basic Books.

Langton, C. G. (1986). Studying Artificial Life with Cellular Automata. *Physica D, Vol. 22*, 120-149.

Locke, J. (1997). *An Essay Concerning Human Understanding.* New York: Penguin Books.

Müller, V. C. (2021). *Ethics of Artificial Intelligence and Robotics (Singularity).* (E. N. Zalta, Ed.) Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/ethics-ai/#Sing

Russell, S., & Norvig, P. (2020). *Artificial Intelligence - A Modern Approach* (4th ed.). Hoboken: Pearson Education Inc.

Sandberg, A., & Bostrom, N. (2008). *Whole Brain Emulation - A Roadmap.* Oxford: Future of Humanity Institute.

Schneider, S. (2019). *Artificial You - AI and the Future of your Mind.* New Jersey: Princeton University Press.

Schrödinger, E. (2012). *What is Life?: With Mind and Matter and Autobiographical Sketches.* Cambridge: Cambridge University Press.

Searle, J. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences*, 417-457.

Searle, J. (1990). Is the brain a digital computer? *Proceedings and Addresses of the American Philosophical Association 64*, (pp. 21-37).

Seth, A. (2021). The Real Problem(s) with Panpsychism. *Journal of Consciousness Studies 28(9)*, 52-64(13).

Sober, E. (1992). Learning from functionalism - Prospects for strong artificial life . *Artificial Life II (Santa Fe Institute Studies in the Sciences of Complexity)* (pp. 749-766). Redwood City: Addison-Wesley Publishing Company.

Tiku, N. (2022, June 11th). The Google engineer who thinks the company's AI has come to life. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

Turing, A. (1950). Computing machinery and intelligence. *Mind*, 433-460.