Challenging cognitivism: computational versus embodied and embedded creativity in AI generative music

Matheus P. Garcia[1*]

[1]Núcleo Interdisciplinar de Comunicação Sonora – Unicamp
R. da Reitoria, 165 – Cidade Universitária, Campinas – SP

m184383@dac.unicamp.br

Abstract. *This paper aims to discuss different perspectives on the nature of cognition in relation to artificial intelligence as a tool for artistic creativity. We bring forth the debate between two important lines of thought in cognitive science: first, we present a brief overview of the computational and representational theory of mind, which posits that the mind functions similarly to a computer, gathering information through the sensory organs and processing this information by means of an internal, symbolic language. We argue that this cognitive point of view is closely linked to the development of the first artificial neural networks and represents a dominant perspective in the field of artificial intelligence to this day. Then, we discuss the idea of embodied and extended cognition, which differs radically from the first by stating that the process of acquiring knowledge is carried out not solely by the brain, but by the entire organism, engaging in constant exchange with its environment through complex mutual modeling. From this debate, we explore the field of generative artificial intelligence by presenting artistic experiences and perspectives in AI-generated music that challenge the cognitivist notion that the brain and computer function in fundamentally similar ways.*

1. Introduction

In 1943, Warren McCulloch and Walter Pitts published a seminal paper that proposed a mathematical model of the neuron and neural networks. This model, influenced by Alan Turing's modern theory of computing and experiments with squid neurons, is considered the first modern mathematical model of the brain [1]. In this model, the neuron functions as a simple mathematical operator, connecting with other neurons via synapses of varying intensity. The neuron sums multiple bioelectric stimuli and, based on this value, decides whether to fire a signal to subsequent neurons through an internal activation function. If the total input signal exceeds the activation threshold, the neuron fires a signal via its output connections to another group of neurons. The McCulloch-Pitts model depicts the neuron as a processing unit with an input interface (synapses) and an output interface (axon). Although individually the neuron processes a basic conditional function (if $x > a$, then 1, else 0), in a network, neurons can form a complex information processing organ [2].

The work of McCulloch and Pitts, as well as others such as Donald Hebb, Jerry Fodor, and David Marr, contributed to what became known as cognitivism and the computational theory of mind (CTM). This theory describes the mind as a computer, not as an analogy, but in a logical-mathematical sense: the brain receives sensory stimuli from the body; these *modal* stimuli are transformed into *amodal* representation and processed; the brain then sends motor commands back to the peripheral body. Thus, the brain functions as a complex system of *inputs* and *outputs* — as, in fact, a computing unit. This line of thought was closely linked to the development of the first artificial neural network systems in the second half of the 20th Century: when Frank Rosenblatt's creation of the Perceptron, the first form of this technology, not only drew from cognitivism's understanding of the brain but also provided empirical validation for some CTM hypotheses [3].

However, the paradigm of the brain as a computer was later highly contested in cognitive studies. The theory of 4E Cognition, derived from the work of authors such as Francisco Varela, Humberto Maturana, Evan Thompson, and Eleanor Rosch, strongly opposes the proposal of a computing brain. This perspective argues that, contrary to the Cartesian dichotomy of body and mind, the cognitive process takes place throughout the entire organism, based on the relationships established with its environment. The authors also contend that the process of constructing knowledge and action are necessarily connected, so that "all doing is knowing, and all knowing is doing" [4].

This paper will explore how the computational theory of mind and artificial neural networks have developed together, in mutual collaboration and validation. We will present some of CTM's hypotheses and considerations about the process of knowing and how it is related to the emergence of machine learning. Additionally, we will present some proposals that contradict CTM, in particular the hypothesis of the embodied mind formulated by Maturana and Varela (2004). We will examine how knowledge arises from the interaction of organism, environment, and social context, with no discontinuity between these aspects. These ideas will be articulated with discussions in the field of artificial intelligence to investigate artistic experiments and perspectives that challenge mainstream ideas on the role of creativity in AI, and the role of AI in creativity. We argue that many of the ideas proposed by CTM have had a strong influence on the current debate on computational creativity and machine learning particularly in the field of generative art, and present divergent artistic perspectives that can be supported by the more contemporary ideas of embodied and embedded cognition. We also draw from the contributions of Vear et al. [5], who created a framework for understanding different forms of human-AI interaction in musicking —

[*] Supported by CAPES

particularly from the first part of their framework, which focuses on how it manifests in human-AI relationship. We will thus be describing aspects of the artistic works mentioned in terms of their *liveness* and *presence* [5].

2. Computational theory of the mind

The computational theory of mind in its modern form was proposed by Hilary Putnam in 1967 and developed further by Jerry Fodor during the 1960s, 70s, and 80s [1]. According to Fodor, the mind shares fundamental aspects of its functioning with computers, although not in the conventional sense of everyday computers, with software made of lines of code and hardware made of silicon. Fodor is mostly influenced by Alan Turing's [1912 – 1954] writings on computer science, in which he describes what became known as the Turing Machine: a hypothetical, abstract computer which, together with circles, triangles, and points, belongs to the abstract realm of mathematics and does not rely on the concrete realization of its workings. It is evident that the brain is very different to computer hardware, but Fodor argues that this difference is superficial when considering the deeper, fundamental similarities between the mind and the theoretical computer.

According to CTM, the brain, like a computer, receives information from the external world in a variety of modalities: visual, sonic, tactile, olfactory, and gustatory sensations, among others. This modal information is transformed into amodal representation: in computers, this means that it is translated into a binary machine language. Similarly, external phenomena exist inside the brain through representations of the external world. This is called the Representational Theory of Mind [3]. It should be noted that Fodor did not attempt to decode or to describe precisely this internal language of the mind. In fact, Fodor states in his book *The Language of Thought* [6] that his commitment to the representational idea was provisional: he argues that the only "remotely plausible" psychological models of the mind at the time were computational [6]. His effort was to then develop a theory with the most scientifically solid ideas available, aware of its possible limitations.

Both computers and minds are, in essence, information manipulators. Their internal states represent external objects, events, places, abstract ideas, and a multitude of things. This information is used to perform motor tasks, solve problems, make plans, or simply stored as memory. In short, both the brain and the computer learn about the world through the symbolic processing of external stimuli (inputs). In response, the brain gives motor commands to the body (outputs), which interacts with the world. We can compare the relationship between mind and body to that between software and hardware.

2.1. Networks of neurons

The neural model proposed by McCulloch and Pitts in 1943 operates as follows:

1) Each neuron is connected to several others at its input interface by means of synapses. Thus, the inputs x_n *are* multiplied by the *corresponding* weights w_n, which are equivalent to the intensity of the synapses.
2) The neuron sums all the inputs, such that

$$h = \sum_{i=1}^{n} w_i x_i$$

3) The neuron decides whether or not to fire based on an activation function $g(h)$, typically a threshold function.

$$g(h) = \begin{cases} 1 \; if \; h > \theta \\ 0 \; if \; h \leq \theta \end{cases}$$

A single neuron is thus capable of performing the basic Boolean operations AND, OR, and NOT, which are fundamental for modern computation. However, for more complex operations such as XOR, a single neuron is not enough — at least two neurons in a network are needed.

The first form of artificial neural network, the Perceptron, was developpped by Frank Rosenblatt in 1958. It consists of McCulloch-Pitts neurons arranged in a network with an input and an output layer of neurons. Machine learning essentially involves of the automatic adjustment of each weight *w* based on mechanisms for evaluating and correcting the network while its structure remains fixed; only the numerical parameters change [7]. Typically, it is necessary to carry out several iterations of a learning cycle, but eventually the network is expected to reach a stable configuration, capable of providing correct answers to a given problem.

This first simple configuration paved the way for more advanced neural network techniques in later decades: multiple layers of neurons were chained together; in the 1980s, the invention of backpropagation was an important innovation that enabled faster and more efficient learning; and more recently, we have seen the combination of multiple complete neural networks in the same architecture, as in the case of adversarial networks. Advanced neural network architectures are beyond the scope of this article; it suffices to know that by combining simple conditional processing units such as the McCulloch-Pitts neuron, it is possible to process information in highly complex ways.

The principle of associative memory is important both for artificial neuron networks and for the computational theory of mind. A neuron, or set of neurons, can be associated with specific ideas, objects, places, etc., so that the sensory stimuli corresponding to these things in the external world have the effect of activating that specific group of neurons. When two groups of neurons, corresponding to different ideas, are frequently activated simultaneously, the synapses between the two are strengthened. Therefore, one mental idea can be strongly associated with another, and the activation of one group tends to trigger the other through reinforced synaptic connections. For example: if every time one visits their grandmother's house she bakes a carrot cake, it won't be long before the idea of a carrot cake and the sensory stimuli linked to it will activate, by association, those neurons linked to the grandmother, because the brain has been conditioned to strengthen the specific synapses that associate one thing with the other in the mind. In neuroscience, this principle was described by Donald Hebb in 1949, in what became known as Hebb's Law. It states that neurons that are frequently activated together will form an increasingly intense synaptic connection [7].

2.2. Artificial intelligence and cognitivism

While McCulloch, Pitts, and Hebb's research on the neuron enabled the invention of the first neural networks, experiments with artificial neural networks also contributed to the development of the computational theory of mind: the study of neural networks in computing provided, in many instances, empirical validation for the cognitivist hypothesis. Artificial neural networks developed in close tandem with CTM, in an interdisciplinary relationship [3]. The field of cognitive sciences emerged in the 1940s with contributions from researchers in psychology, neurophysiology, engineering, mathematics, and anthropology [8] in what became known as cybernetics, or the study of the mechanisms responsible for self-controlled or self-regulated behaviors in machines and animals.

With the advent of backpropagation in the 1980s, machine learning with neural networks demonstrated how a computer might simulate vision, complex symbolic pattern recognition and language acquisition based on the cognitivist theory. In 1986, David Rumelhart and James McClelland proposed a model capable of taking English verbs as input and returning their simple past forms as output. Their work departed somewhat from classical cognitivism by stating that knowledge occurs in a *distributed* and *parallel* manner in the brain — much like how knowledge in an artificial neural network model is not stored in conventional memory space, but distributed across the different values of each neuron and synapse (weights).

3. Different perspectives on cognition

John Searle, a prominent critic of CTM, challenged the cognitivist theory with what became known as the Chinese room argument. In this anecdote, a person is locked in a room, and his only form of communication is through letters written in Mandarin, to which he must reply. This person knows nothing about the Chinese language, but they have a detailed instruction book that demonstrates how to provide coherent answers in Mandarin algorithmically. Despite succeeding in their task, the prisoner doesn't know how to speak Mandarin, but merely to simulate knowledge of the language. For Searle, this is exactly what computers do: they acquire information and transform it through formal rules into output, with no real understanding of the symbols they manipulate [9].

Among the theories critical of cognitivism, we are particularly interested in what has become known as 4E Cognition, derived from the framework of what Varela and Maturana [10] called *autopoiesis*. The four E's refer to different, but closely related and overlapping aspects of cognition and creativity: *embodied*, *embedded*, *enacted*, and *extended*. In this section, we will provide a brief overview of each of these four dimensions of cognitive processes and how they relate to musical creativity, while demonstrating how these ideas are present in the artistic-scientific work of artists and researchers in sound art. For this, we take the scientific contribution of Van Der Schyff et al. [11] in developing a framework for 4E Cognition as applied to musical creativity.

3.1. Embodied and embedded

While classical computationalism provided important theoretical insights for cognitive sciences, some problems and limits of this theory must be acknowledged. Through the lens of CTM, many aspects of doing and knowing remain without convincing explanation: while it is possible to simulate computationally many processes involved in human and animal behavior, other fundamental tasks of life remain major challenges for a computer. What became known as Moravec's paradox is the statement that tasks that are easy for humans are difficult for computers, and tasks that are difficult for us are easy for machines [12]. In a context where the domain of possibilities is finite and well defined, computers tend to be quite successful. For example, in games such as chess or Go, computers can easily calculate thousands or millions of possible scenarios following each move. By the 1990s, chess was effectively solved: computer programs could defeat even the most skilled grandmasters.

However, there are things very simple for a human being or an animal and notoriously difficult for a computer, even today: balancing on two legs; moving through an uncontrolled environment; interpreting emotional states through body expressions; understanding colloquial speech; or grasping the simplest abstract philosophical concepts. This is because, in these situations, it is impossible to define a finite number of variables. The extent of the problem is ultimately as big as the world itself, in all its chaotic unpredictability. While the knowledge acquired by the machine is limited by the finite set of information provided to it, part of human knowledge was shaped over billions of years of evolution, as our ancestors needed to understand the world around them and survive in it. From the 1970s onwards, it became clear that, contrary to cognitivist assumptions, simulating simple cognitive processes with a computer "requires a seemingly infinite amount of knowledge, which we take for granted [...] but which must be spoon-fed to the computer" [13].

To explore the *embodied* aspect of cognition, let us take the example of the cricket, an insect with very limited cerebral capabilities when compared to humans. The female cricket in its natural habitat needs to identify the sound of the male cricket in order to move towards him when she wants to mate. How does she do this? One might think from the cognitivist point of view that the cricket's ears send a coded signal to the brain, which then processes it, identifies the sound of interest within a noisy environment, determines its direction, and finally sends a command to the motor system. However, this is not the case: the cricket's solution is to listen with its legs. Its sensory organ is located in the motor system, closely connected to it, so that the ear (or, the leg) that first receives the sound stimulus also receives a direct order to move, without the need for any central computational mechanism [3]. The theory of embodied cognition states that, for human beings alike, gathering information, processing it, and responding to it, all take place throughout the body. To support this, Varela [14] notes that the neural connections between the central nervous system and sensory-motor system in mammals is not a linear one-way flow of information. Rather, the dynamic underlying the perceptual-motor task is that of a complex two-way, highly cooperative system in which information is constantly exchanged between the brain and the sensory and motor system.

Proponents of embodied cognition reject the idea that the brain is a central processor, and that creative behavior is the result of internal and abstract mental models [3]. According to this theory, creativity emerges more directly from perception–action coupling, without relying on representational models as mediators. Since cognition is the result of an evolutionary process, embodied cognition considers that sensory perception, motor action, and creativity have been constituted together in the environment that human beings have inhabited over time. This gives cognition its *embedded* character: "There is no discontinuity between what is social and what is human and their biological roots. The phenomenon of knowing is all of one piece, and in its full scope it has one same background" [4].

3.2. Enacted and extended

Let's consider the question: what came first, the world or the image of the world? There are two opposing perspectives to answer this. The first states that the human mind interprets the world through vision, and the world exists externally independently whether it is perceived. The second considers that the world is a mental projection of the cognitive system itself and its internal laws. Embodied cognition stands between these two perspectives [3].

Knowledge is not simply information existing in the external world, waiting to be passively observed. Rather, knowledge is obtained through the mutual shaping of a shared world. It is dependent on the actions of the creative subject, and through this effort both the environment and organism are transformed. One cannot exclude themselves from the world in order to perceive it — we are immersed in it.

The categories of perception, thus, belong to a shared world that is simultaneously biological and cultural. The world and the person who perceives it mutually specify each other. This reasoning also implies that action and perception are inseparable — it is in the process of creating that knowledge emerges, and every motor action is embedded in a biological, psychological, and social context. According to the perspective of embodied cognition, therefore, there is no external world that can be retrieved by the senses, but rather "common principles or regulated links between the sensory and motor systems that explain how action can be perceptually oriented in a world dependent on the observer" [13]. The environment does not exist in a separate system from our perception-action — the latter also contributes to and is part of the environment. From this idea, Varela and Maturana coined the word *enaction*: action guided by perception, or learning as an active creative process carried out by an autonomous subject in relation to the environment [4].

An example of how we can empirically verify the enacted nature of knowledge is the experiment carried out with kittens by Held and Hein [3]. In this experiment, a group of cats was raised in the dark and only exposed to light in a controlled situation. Among these cats, one group was allowed to walk normally, and each individual was attached to a trolley that carried a cat from the other group, and depended on the first to move through space. When freed and allowed to move on their own, the second group, entirely passive in the first stage of the experiment, behaved as if they were blind, bumping into objects and falling down steps. The visual experience, which was shared by both groups, has no learning effect unless it is combined with the motor experience, which only one of the groups was allowed to have.

Finally, the *extended* aspect of cognition is an outcome of the *embedded*. It refers to the specific roles that the tools, such as musical instruments or computers, can have in creativity. If creativity takes place inside and outside the body of an individual, then we must admit that the interaction between the individual and the objects closely involved in an activity is integral to the process of knowing. If a musician has a particularly close connection to a particular instrument, it seems that a big part of his musical knowledge can only be easily accessed when this interaction takes place.

3.3. Modality

4E cognition fundamentally disagrees with the idea that the mind and cognitive processes happen in a representational way. For cognitivism, knowing involves transforming sensory information into a symbolic internal language that can be easily translated across different senses due to its amodal character. In contrast, embodied cognition posits that knowledge does not occur independently of the senses, and its modality is intrinsic to the knowledge itself. This does not mean that we are incapable of translating knowledge into different modalities, such as abstracting a tactile sensation into a sound one in a comprehensible way. In musical practice, the modality of knowledge can be easily observed: so many players will rely on muscle memory in order to recall the notes of a song, even if that means moving their fingers in the air as though the instrument was present. Much of a musician's knowledge is intrinsic to their interaction with the tools (instruments) involved in their practice and the sensory feedback – not only sonic, but also tactile – they receive while performing.

Knowledge does not occur in a representational way because it corresponds to the internal state of the nervous system at each interaction with the environment. The nervous system, in turn, was not designed as a computer, but is the result of phylogenetic drift, and is best described as "a unit defined by its internal relations in which interactions come into play only by modulating its structural dynamics, i. e., as a unity with operational closeness" [4].

Watanabe [3] offers empirical evidence of the modal nature of knowledge. In an experiment, subjects observed pairs of three-dimensional geometric figures and determined whether the two images corresponded to the same object as seen from two different angles, or whether they were different shaped objects. In this experiment, it was observed that the participants' response time was directly proportional to the difference in the angle of rotation between the pair of figures in the cases where the objects were identical in shape: the greater the angular difference between the pieces, the longer it took the participants to identify their equivalence. This indicates that, mentally, the participants simulated the action of rotating the object in order to reach a conclusion — the sensory-motor mechanism is was engaged, even if the motor action itself is not required. For a computer, on the other hand, it would be enough to translate the image into symbolic language and then carry out the necessary calculations, and in this way the decision-making time would be equivalent regardless of the angle of the figures.

The conclusion is that the use of a concept is a sensory-motor simulation that involves brain and body, and which refers to situations we have experienced in the past. Since these experiences emerge from the complex interaction between the body, the environment and the social context, the modality of this knowledge is understandable.

3.4. Artificial intelligence and the problem of knowledge

We have seen how, since the 1950s, the science of the brain and cognition has developed in an interdisciplinary way: computing and neuropsychology have contributed to each other on a common knowledge base. In the same way, we can establish links between the debate on cognition, which we have just presented, and current debates on artificial intelligence.

The 2010s saw the emergence of a new "spring" in artificial intelligence with the advent of deep learning (Mitchell, 2021). In recent years, we have seen a dizzying advance in the capacity of neural network models in areas such as **n**atural language, image, video, and sound modeling (Gozalo-Brizuela; Garrido-Merchan, 2023). This period of renewed novelty in artificial intelligence raises multiple debates about the possibilities posed by this technology in the 21st century. On the one hand, we have an optimistic perspective on artificial intelligence, which argues in favor of a continuous movement towards genuine artificial intelligence. In contrast, there are those who stand for a more modest analysis of recent models and innovations as capable of responding with unprecedented efficiency to a still limited series of tasks. We believe that the debate between the computational theory of mind and embodied cognition is quite pertinent when considering the following question: are neural network models actually moving towards intelligence comparable to that of a human?

A much-debated article on this question was published by Bender et al. [15]. The authors argue that the observed performance of large language models (LLMs) is often confused, even by computer experts, with an actual understanding of language. Language models, however, are nothing more than tools capable of finding the most *statistically probable* alternative to a proposition, which is achieved by training on enormous datasets. The authors compare the behavior of large language models to "stochastic parrots": they are only capable of making a probable recombination of the database that constitutes their universe. Like a man who learns to answer letters in Chinese without understanding a word of what he writes, language models do not, by their very nature, have the ability to actually understand the symbols they manipulate - only to simulate knowledge of the language. The authors' concern lies precisely in models that reproduce the hegemonic worldviews of the databases on which they are trained and the arbitrary taxonomic choices of the developers who created these databases, without being able to identify, for example, hate speech that is possibly present in this *corpus* [15].

The author Melanie Mitchell [16], in her article "*Why AI is harder than we think*", describes some of the fallacies that she believes undermine the inflated discourse surrounding recent innovations in the field of machine learning. The first point she makes is that it is wrong to interpret a given technological advance as a step forward in a linear and continuous trajectory. For her, the positive performance of a model within what she calls "*narrow intelligence*" has nothing to do with "*general* intelligence". In other words, we can observe the performance of a chimpanzee that manages to climb higher and higher trees more efficiently. However, this in no way brings the chimp closer to landing on the moon. The ability of language models to make progress within a well-defined problem, on a finite information base, is not along a continuous line with actual general intelligence: as we discussed earlier, the latter needs to be able to deal with an infinite number of variables, in a universe where problems are not delimited by a restricted space of possibilities like a game of chess. 4E cognition offers a vastly more complete and complex framework for understanding the process of acquiring knowledge that goes well beyond the capabilities of a computer: it involves the organic interaction with one's environment and peers. Much more than to gather information, to know is to act upon the world, to make use of tools, to collaborate and to transform and be transformed by the world. It is inseparable from one's emotions, moral values, and subject perceptions. Also, knowledge it is not only the process of one individual through the course of a lifetime, but has been built by living beings since the first organic molecules appeared some 2 billion years ago [10].

The author is also in agreement with further claims by Varela, Thompson and Rosch [13] when she argues that very easy problems for human beings are notoriously difficult for computers, and that easy problems for computers can be impossible for human beings, such as performing thousands of mathematical operations in a fraction of a second. She also draws upon insights from the field of cognition by stating that human intelligence, unlike that of computers, is not located in a central processing organ but is distributed throughout the body [16].

According to some respected artificial intelligence researchers, she says, the problem of artificial intelligence will be solved once we can match the processing power of the brain's wetware through computer hardware. After all, the largest artificial neural networks only have billions of connections, while the brain has trillions. However, the debate presented above makes this a difficult claim to defend plausibly. The idea of a super-intelligent computer with pure rationality escapes the considerations of psychology, which states that not only is intelligence necessarily linked to corporeality, but also that reason does not occur independently of the emotional and social.

3.5. 4E musical creativity

So far, we have discussed some relevant aspects of Varela's (and some of his colleagues') work on cognition, mainly in respect to how knowledge is acquired by human and other living beings through "mutual modeling", which occurs in the exchange between individuals and their environment in an embodied form. We have highlighted the main points of diversion from this theory of embodied and embedded cognition to previous works by cognitivist researchers such as Fodor [6], especially in how — and if at all — information is gathered and processed by the organism.

In this chapter, we will explore how some ideas from 4E cognition can and have become a conceptual basis in sound art. Specifically, we will examine two composers and researchers influenced by Varela's ideas, who challenge dominant cognitivist views in their approach to using artificial intelligence for music creation. A notable example is Holly Herndon's project PROTO, which explores deep learning technology for sound composition and synthesis.

4. Challenging cognitivism

PROTO is the name of a music ensemble and album released in 2019 by Holly Herndon in which she explores deep learning technology for composing and synthesizing sounds. Together with Matt Dryhurst, creative partner, and Jules LaPlace, software developer, Herndon created the program *Spawn*, a neural network system that was used in her album in various ways. In this work, Herndon aims to develop the particular sonorities recently made possible by neural synthesis, in a process that she sees as a continuation of the exploration of sonorities that took place in the 20th century. Furthermore, with Spawn, she demonstrates artificial intelligence without exaggerating its capabilities, while highlighting future paths for musicians interested in AI in a certain optimistic manner.

The first aspect of Herndon's approach to AI in PROTO that relates to an *embedded* concept of creativity is precisely its collectiveness. The development of the deep learning program relies on her work with a vocal ensemble, in which *Spawn* and PROTO are mutually tied in musical practice and composition, one system informing and influencing the other. Herndon describes how the performance of the ensemble is recorded and used as training set for *Spawn*'s machine learning, while the preliminary results from this first stage of development are also used as material and source of inspiration for the choir. There is a feedback loop from the ensemble and the program, each adding a particular layer of sound complexity:

> Our process consisted of my writing a vocal idea either as a score or as an audio file, which the ensemble would perform and I would record. I would then process the recorded files in the studio and return with modifications to the score. Sometimes the most interesting developments came from the ensemble emulating the processing itself, or through experiments in hierarchies and roles within the group. [17]

Processing the ensemble recordings through a neural network presents particular characteristics, like what she describes as a "nasal quality" to the voice, which the ensemble later incorporated in performance. This led the group to search inspiration in some genres of folk music from mountainous regions in Europe [17]. In this sense, even though PROTO is not *live* in musical practice as it has no real-time performance capabilities, its *presence* is meant to be that of a co-creative partner in musicking [5].

As an academic writer, Herndon emphasizes that her vision of AI in music is, or should be, of an embedded form of digital creativity. The reader should be aware that the rise of deep learning models in the last decade has spawned numerous debates on intellectual property, copyright and the nature of human versus computer creativity [15], [18], [19], [20], [21], [22], [23], [24].

Aware of this background discussion, Herndon's work defies the idea that neural networks are capable of artistic creation without human effort. On the contrary, *Spawn* is the product of a great amount of human work.

Herndon's approach to embedded creativity was influenced by George Lewis's ideas on human-software improvisational interaction in African-American music [25]. He was a pioneer in his work with the program *Voyager*, an interactive computer music system which he described as a "virtual improvising orchestra" [25].

> Musical computer programs […] represent the particular ideas of their creators. As notions about the nature and function of music become embedded into the structure of software-based musical systems and compositions, interactions with these systems tend to reveal characteristics of the community of thought and culture that produced them. Thus, it would be useful here to examine the implications of the experience of programming and performing

with Voyager as a kind of computer music-making embodying African-American cultural practice. [25]

In Lewis' Voyager, the communal aspect of his system is mostly manifested in aesthetic form. He explores the improvisational nature of black music in North America in what he calls "too many notes": a tendency towards a high density of events in a short time frame, a sustained use of "rapid, many-noted intensity structures" that can be observed in the performances of African-American improvisers such as Cecil Taylor, John Coltrane and Albert Ayler [25]. In addition, his intention is to dissolve the commonly found hierarchy of human performers over computer systems in music performance is reminiscent to African musical traditions in which instruments are elevated to an almost human status. The *presence* and *liveness* of this AI is meant to be as great as possible, as it interacts with the improvising musician in real time as a co-creative agent.

For Herndon and Lewis, creativity is not a matter of expanding one's knowledge of the mind through computational models. On the contrary: their exploration is by attempting to elevate a computer system to human forms of creativity. Herndon goes to the extent of referring to her system as a "child […] raised by a community" [17]. In her words, PROTO, the music ensemble, and Spawn, the computer system, feed off each other's particular qualities in mimetic interaction: the "pre-rational, or not-yet-rationalized, mode of behavior, with an affinity towards the sensuous and embodied, non-conceptual re-enactment of cognitive processes" [17]. She makes reference to Christopher Small's concept of "musicking": musical practice as a set of relationships that goes beyond the agents traditionally thought of as musicians, but comprehends a larger and more complex social phenomenon. Musical interaction occurs between the performer and their tools; in between performers; and ultimately between performers and the social complex and its ever-changing cultural and political structures. This goes very much in agreement to Varela's idea of embedded cognition, as we have previously discussed.

Herndon's option for a computer system that directly processes digital audio (instead of a symbolic one, that operates over MIDI encoding, for example) can not only be attributed to the influence of 20th century music of sound masses, but is also a way to further explore the embodied and embedded aspect of music making and creativity. The voice sounds that feed *Spawn*'s neural model and the voice-like sounds it puts out represent the agency of human individual performers and its collective form, the choir. This "maintains traces of the human bodies and performances that created the training sets"[17]. The results obtained from preliminary neural training can also be used as sound material to further Herndon's embodied exploration of her and her ensemble's voices.

5. Conclusion

In this paper, we discussed how different perspectives on human cognition and creativity can relate to the way we make and the way we interpret artificial intelligence.

It should be acknowledge that the artistic works discussed in this paper have a limitation in the sense that they do not attempt to create any form of *machine embodiment* in an actual sense — as has been attempted with the use of robotics, for example, in the work by Vear and Benerradi [26]. Instead, the forms of embodiment that could be observed can only be present in the datasets used to create the models and software used by Lewis and Herndon.

Saint-Germier, Canonne and Fiorini [27], in their work with the software Somax2, have demonstrated ways in which the embodiment of music performance, even in the limiting form of symbolic representation (MIDI protocol) used as training data for a model, can have an effect in human-machine musical interaction in the context of improvisation. We also believe that the discussion of 4E cognition regarding AI models in musicking should consider embodiment music in regards not only to music performance in a bio-mechanical sense, but in regards to all forms of musical practice including creativity and listening.

Neural networks and other forms of AI are an ever-growing part of the way we make art in the 21st Century. In scientific research, many are asking the question of what can AI do for artists and how it could change the way artists, musicians, and writers work in their craft. However, while scientific breakthroughs have had lasting impacts on art throughout history, research and experimentations in artistic practice can also have an impact on scientific knowledge. Engaging art can often lead to novel points of view on the subjects of our everyday life and can bring forth radical new ways of understanding the world around us. Because of this, we believe that in addition to asking what can science do for art, an equally important concern is *what can art do for science?*

Acknowledgments

Much of this paper was originally written as a requirement for the course on music cognition taught by Prof. Dr. Manuel Silveira Falleiros, Prof. Dr. José Fornari, and Prof. Dr. Vilson Zattera at the Núcleo Interdisciplinar de Comunicação Sonora (NICS) at Unicamp. Thanks to these professors, I had access to many of the authors mentioned here. I would also like to thank my colleague Lucas Quínamo for introducing me to the writings of Francisco Varela through his masters dissertation [28].

References

[1] G. Piccinini, "The First Computational Theory of Mind and Brain: A Close Look at Mcculloch and Pitts's 'Logical Calculus of Ideas Immanent in Nervous Activity,'" *Synthese*, vol. 141, no. 2, pp. 175–215, Aug. 2004, doi: 10.1023/B:SYNT.0000043018.52445.3e.

[2] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, Dec. 1943, doi: 10.1007/BF02478259.

[3] C. H. A. Watanabe, "A teoria computacional da mente

e o dilema de Searle," Mestre em Filosofia, Universidade Estadual de Campinas, Campinas, 2021. doi: 10.47749/T/UNICAMP.2021.1235358.
[4] H. R. Maturana and F. J. Varela, *A Árvore do Conhecimento*. São Paulo: Palas Athena, 2004.
[5] C. Vear, S. Benford, J. M. Avila, and S. Moroz, "Human-AI Musicking: A Framework for Designing AI for Music Co-creativity," 2023.
[6] J. A. Fodor, *The language of thought*. in The Language & thought series. New York: Crowell, 1975.
[7] S. Marsland, *Machine Learning: An Algorithmic Perspective*, 2nd ed. Chapman and Hall/CRC, 2014. doi: 10.1201/b17476.
[8] L. F. Oliveira, "O DEBATE SOBRE O REPRESENTACIONALISMO NAS CIÊNCIAS COGNITIVAS," *Kínesis*, vol. 8, no. 17, pp. 85–114, Nov. 2016, doi: 10.36311/1984-8900.2016.v8.n17.06.p85.
[9] J. Searle, *Minds, brains and science*. Cambridge: Harvard University Press, 1984.
[10] H. R. Maturana and F. J. Varela, *The tree of knowledge: The biological roots of human understanding*. Boston: Shambhala, 1992.
[11] D. Van Der Schyff, A. Schiavio, A. Walton, V. Velardo, and A. Chemero, "Musical creativity and the embodied mind: Exploring the possibilities of 4E cognition and dynamical systems theory," *Music & Science*, vol. 1, p. 2059204318792319, Jan. 2018, doi: 10.1177/2059204318792319.
[12] F. Truck and H. Moravec, "Mind Children: The Future of Robot and Human Intelligence," *Leonardo*, vol. 24, no. 2, p. 242, 1991, doi: 10.2307/1575314.
[13] F. J. Varela, E. Thompson, and E. Rosch, *A mente incorporada: Ciências cognitivas e a experiência humana*. Porto Alegre: Artmed, 2003.
[14] F. J. Varela, *Ethical know-how*. Stanford: Stanford University Press, 1999.
[15] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Canadá: ACM, Mar. 2021, pp. 610–623. doi: 10.1145/3442188.3445922.
[16] M. Mitchell, "Why AI is Harder Than We Think," Apr. 28, 2021, *arXiv*: arXiv:2104.12871. doi: 10.48550/arXiv.2104.12871.
[17] H. Herndon, "Proto," Stanford, Califórnia, 2019.
[18] K. Crawford and T. Paglen, "Excavating AI: the politics of images in machine learning training sets," *AI & Soc*, vol. 36, no. 4, pp. 1105–1116, Dec. 2021, doi: 10.1007/s00146-021-01162-8.
[19] T. Silva, "Colonialidade difusa no aprendizado de máquina: camadas de opacidade algorítmica na ImageNet," in *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*, São Paulo: Autonomia Literária, 2021, pp. 84–104.
[20] J. ZYLINSKA, *AI Art: Machine Visions and Warped Dreams*. in MEDIA : ART : WRITE : NOW. Londres: Open Humanities Press, 2020.
[21] H. H. Jiang *et al.*, "AI Art and its Impact on Artists," presented at the AIES '23: AAAI/ACM Conference on AI, Ethics, and Society, Montréal: ACM, Aug. 2023, pp. 363–374. doi: 10.1145/3600211.3604681.
[22] T. Broad, F. F. Leymarie, and M. Grierson, "Network Bending: Expressive Manipulation of Deep Generative Models," Mar. 12, 2021, *arXiv*: arXiv:2005.12420. Accessed: Oct. 11, 2024. [Online]. Available: http://arxiv.org/abs/2005.12420
[23] A. Baio, "Invasive Diffusion: How one unwilling illustrator found herself turned into an AI model," Waxy.org. Accessed: Oct. 18, 2024. [Online]. Available: https://waxy.org/2022/11/invasive-diffusion-how-one-unwilling-illustrator-found-herself-turned-into-an-ai-model/
[24] M. Tromble, "Ask not what AI can do for art... but what art can do for AI," *Artn.*, no. 26, Jul. 2020, doi: 10.7238/a.v0i26.3368.
[25] G. E. Lewis, "Too Many Notes: Computers, Complexity and Culture in 'Voyager,'" *Leonardo Music Journal*, vol. 10, pp. 33–39, 2000, [Online]. Available: http://www.jstor.org/stable/1513376
[26] C. Vear and J. Benerradi, "Jess+: designing embodied AI for interactive music-making," Dec. 09, 2024, *arXiv*: arXiv:2412.06469. doi: 10.48550/arXiv.2412.06469.
[27] P. Saint-Germier, C. Canonne, and M. Fiorini, "The corpus' body. Embodied Interaction from Machine-Learning in Human-Machine Improvisation.," 2024.
[28] L. Quínamo, "Entre a imagem, a prática e a música: analisando os processos criativos de A Dream Within A Dream, de Tatiana Catanzaro, e de Furnas 1, de Felipe de Almeida Ribeiro," Dissertação, Campinas, 2023. [Online]. Available: https://www.repositorio.unicamp.br/acervo/detalhe/1384840