A Digital Touch[1]

The »Body Issue« in Computational Creativity

Mattia Merlini and Stefano Maria Nicoletti

Introduction

Understanding if machines will ever be as creative as humans is not only an interesting question from a technical standpoint, but it is also relevant from a philosophical perspective as it helps us to dive deeper into what defines human experiences and human nature itself. As a testimony of the relevance of the question, some of the brightest minds in computer science and philosophy have put different opinions and diverse technologies to the test against the creative processes needed to produce literature, poetry, visual arts and – of course – music[2], from the era of Charles Babbage[3] and Ada Lovelace, through the ideas of Alan Turing, until the present day.

In spite of the existence of promising results and methodologies in the field of Computational Creativity – and in the category of *Music Generation Systems* in particular[4] – we are convinced that there are at least four reasons – or »is-

1 Stefano M. Nicoletti's work has been funded by the ERC Consolidator grant 864075 CAESAR.

2 See Harold Cohen, »The Further Exploits of AARON, Painter«, in: *Stanford Humanities Review* 4/2 (1995), pp. 141–158; Simon Colton, Jacob Goodwin and Tony Veale, »Full-FACE Poetry Generation«, presented at: *International Conference on Computational Creativity 2012*, University of Dublin 2012.

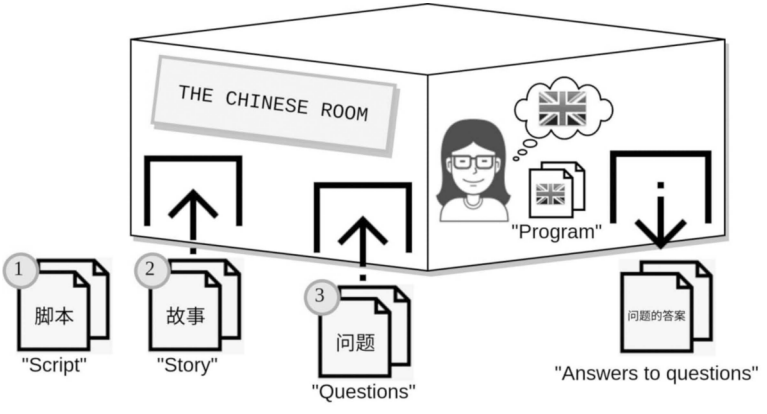
3 See Henry P. Babbage (Ed.), *Babbage's Calculating Engines: Being a Collection of Papers Relating to Them, Their History and Construction*, Cambridge: Cambridge University Press 1889.

4 See Mattia Merlini and Stefano Maria Nicoletti, »Of Flesh and Steel: Computational Creativity in Music and the Body Issue«, in: *INSAM: Journal of Contemporary Music, Art and Technology* 4 (2020), pp. 24–42, here pp. 26–30. This paper is largely based on the research made thereby, while also presenting some notable elements of novelty

sues« as we call them – that keep machines from being (and eventually becoming) truly creative agents. In this paper, we are going to briefly argue against the possibility of using computers and algorithms to *create* a mind which is able to *understand* and to have mental states (the thesis of »Strong Artificial Intelligence«) in order to vouch for the existence of issues which we believe are more relevant and more imminent in dealing with artificial agents in creative contexts. Second, we will sketch these issues to then focus on the one we call »the body issue«, which considers the role of the human body in the experience and creation of music, drawing inspiration from the philosophical positions of Maurice Merleau-Ponty and Roland Barthes, but also from the neuropsychological research on mirror neurons and embodied cognition.

The Chinese Room and Strong Computational Creativity

Figure: The Chinese Room argument's structure.



in case studies analysis and focus on theoretical premises. We thank the editors of *INSAM* for letting us publish this updated and revised version in these proceedings.

First, let us sketch the famous thought experiment of the »Chinese Room« by philosopher John Searle (represented in the figure above).[5] Suppose that I am a native speaker of English and that I know nothing about Chinese. Furthermore, imagine that I am locked in a room and that I am given two batches of Chinese symbols along with a set of rules: the rules are written in English and can be used to correlate the second batch with the first, only by identifying the shapes of symbols on them. Moreover, suppose that I am given a third set of symbols in Chinese. With this, I am given instructions that allow me to correlate the symbols from the third batch to the previous two batches. The people that handed me those three sets of symbols call the first batch a »script«, the second a »story« and the third »questions«. The symbols that I return in response to the third batch are called »answers to the questions« and the set of rules that I am given is the »program«. Now, suppose that I get extremely good at manipulating Chinese symbols and that the programmers are outstandingly capable at writing the program for me. In this case, my responses would be indistinguishable from the ones that a native speaker of Chinese would give and my performance would be equivalent – from the outside – to the one I would give if I was playing the game with English batches. However, despite this excellent performance, I still do not understand Chinese. As Searle puts it:

> »[I]n the Chinese case, unlike the English case, I produce the answers by manipulating uninterpreted formal symbols. As far as the Chinese is concerned, I simply behave like a computer; I perform computational operations on formally specified elements. For the purposes of the Chinese, I am simply an instantiation of the computer program«.[6]

The English and the Chinese versions of this experiment clearly hold a difference for Searle: in the first scenario, I understand the stories because I understand the English language, while in the second, I am only manipulating symbols. If that is the case, and if I am indeed acting like a computational device, we could argue that a computer would not be capable of creating a mind that truly understands and that holds mental states (the thesis of »Strong Artificial Intelligence«). If we accept the argument behind Searle's experiment, we will come to the conclusion that computers are not able to grasp the *meaning* of symbols they manipulate, which is an essential characteristic of a mind that

5 See John R. Searle, »Minds, Brains, and Programs«, in: *Behavioral and Brain Sciences* 3/3 (1980), pp. 417–457.

6 Ibid., p. 418.

understands. Rather, we would be more inclined to adopt the thesis of »Weak Artificial Intelligence« which posits that computational devices provide us with nothing more than powerful tools to understand our hypotheses and study the mind. For the theorists of the »Weak AI« thesis, machines can at best simulate or reproduce intelligent human behaviours, rather than instantiate an understanding mind. If this thesis seems to be convincing, then we could accept the specific distinction between »Strong and Weak Computational Creativity« made by al-Rifaie and Bishop that stems from the aforementioned argument:

> »An analogy could be drawn to computational creativity, extending the notion of weak AI to ›weak computational creativity‹, which does not go beyond exploring the simulation of human creativity; emphasising that genuine autonomy and genuine understanding are not the main issues in conceptualising weak computationally creative systems. Conversely in ›strong computational creativity‹, the expectation is that the machine should be autonomous, creative, have ›genuine understanding‹ and other cognitive states. [...] We fundamentally suggest that Searle's famous thought experiment also targets the analogous notion of ›strong computational creativity‹. I.e. Searle using a similar »room« could get so good at following the rules that the strings of symbols he outputs from the room successfully control a ›Strong‹ computationally controlled creative art-system, producing works judged to have artistic merit by people outside the room, even though Searle-in-the-room remains ignorant of the produced art and the externally labelled ›art practice‹«.[7]

One Problem, Four Issues

If Searle's argument at least convinces us that the advent of truly intelligent artificial agents (i.e., Strong AIs) is not probable nor imminent – which is all we ask in our case – then we are also likely to think that *strong creative agents* are not in sight, following al-Rifaie and Bishop's considerations. In doing so, we can direct our attention on issues which we deem more imminent when talking about computational creativity and music in particular (even when we

7 Mohammad M. al-Rifaie and Mark Bishop, »Weak and Strong Computational Creativity«, in: *Computational Creativity Research: Towards Creative Machines*, ed. by Tarek R. Besold, Marco Schorlemmer and Alan Smaill, Paris: Atlantis Press 2015, pp. 37–49, here pp. 45 f.

consider weak computational agents). These are what we call the social issue, the body issue, the experiential issue and the consciousness issue.[8]

The *social issue* focuses on the social aspect of music (but the same could be said of any art), which is the fact that music is always situated within a social context, responding to certain necessities, social functions and sometimes also ideological positions. Despite the outstanding compositional abilities of AIs, computers are not part of any society (and, surely, not of human society), so their creations bypass this important aspect of creativity. We already have machines capable of creating songs »in the style of the Beatles« (like Pachet's Flow Machines),[9] but of course the meaning of a cultural work goes far beyond the pure phenomenon, which is bound to be »empty« if deprived of a link with a social reality. Music is always a reason for aggregating people, for sharing a passion and maybe even some aesthetic or political ideology. What is so special about a plain and simple ballad like »Imagine« by John Lennon if not the *situated* social meaning of the song, composed in a certain moment by *that* very artist with a specific background, resonating with certain categories of people, and so on?[10] All these features trespass the limits of the pure musical phenomenon, and must be kept in mind when evaluating and understanding the meaning and relevance of a piece of music. There is a socio-semiotic level of understanding music that should not be underestimated.

The *experiential issue* is not too different, but it focuses on the individual experience more than the social issue does: every aesthetic choice, every desire for communicating something via artistic activity, every form of expression by an author (even those that try to eliminate the author's figure from the creative game, as it happens in much structuralist and aleatoric music) always takes place *in a certain way* because of the composer's personal experiences and beliefs, positions and ideas, values and interests; all things that come from the

8 The first presentation of this composite idea can be found in Mattia Merlini and Stefano Maria Nicoletti, »Inhuman, All Too Inhuman: Intrinsic Limits of Computational Creativity in Music«, in: *Riffs: Experimental Writing on Popular Music* 4/1 (2020), pp. 28–46. This was our first attempt at presenting the problem, largely focusing on a general framework including the four issues.

9 See Fiammetta Ghedini, François Pachet and Pierre Roy, »Creating Music and Texts with Flow Machines«, in: *Multidisciplinary Contributions to the Science of Creative Thinking*, ed. by Giovanni Emanuele Corazza and Sergio Agnoli, Singapore: Springer 2016, pp. 325–343.

10 See Lucio Spaziante, *Sociosemiotica del pop*, Roma: Carocci 2007, p. 33.

most »human« side of their existence.[11] We *can* consider the music composed by Iannis Xenakis as a pure musical phenomenon – of course we can! But it is impossible to understand it without taking other elements into consideration as well. Even this music, apparently so easily reproducible for a computer (since it is based on calculation and formulas, trying to emancipate the sonic phenomenon from the control of the human author), is deeply grounded in experiences, ideological and aesthetic positions, idiosyncrasies and insights. Even when reading how Xenakis presented his music, it is evident that his cultural background (his musical education, his Greek identity, his passion for ancient philosophy) and his position in the broader field of musical production of the time, play a fundamental role in the understanding of his opus.[12] When asked about the nature of his music, even the »hyper-rationalist« composer Xenakis would state that no music can come from pure calculation, since the very way human beings are made asks for aesthetic choices that are, so to say, filtered by our peculiar way of »feeling« music.[13]

The *consciousness* issue is even more general and addresses the importance of consciousness in our »trade« with the world, and in particular of knowing the world in a way that is essentially *qualitative*, and not just quantitative as it happens with computers. Thanks to consciousness, there is a direct link with the world. This last point resonates not only with the aforementioned Chinese room argument, but also with positions coming from the phenomenological tradition so apparently distant from Searle's philosophical background. To start with Searle, in his famous argument there is a distinction at play: that between syntax and semantics. The English-speaking person manipulating symbols in the room can produce flawless results from a syntactic perspective – that is, pertaining to the level of symbols and the visual aspects regulating their mutual relationships. Yet she makes no progress when it comes to semantics – that is the level of language pertaining to the meaning of the words. This is because it is impossible to *understand* what words are

11 See Jean-Jacques Nattiez, »Alcuni concetti fondamentali di storiografia della musica«, in: *Rivista di analisi e teoria musicale* 13/1 (2007), pp. 7–35.

12 See Iannis Xenakis, *Musica. Architettura*, trans. by Letizia Lionello, Giancarlo Secco and Angelo Varese, Milano: Spirali 1982 (Orig.: *Musique. Architecture*, Paris: Casterman 1976).

13 See Iannis Xenakis, interviewed by Enzo Restagno, in: *Xenakis*, ed. by Enzo Restagno, Torino: EDT 1988, p. 31.

referring to without having contact with the world or without touching, feeling, needing and »living« a world inhabited by things that are meaningful for us. As an analytic philosopher, Searle probably would not so easily endorse such a phenomenologically (and thus continentally) connotated claim, but at least it brings us to the other big point of reference we ought to mention here, which is, of course, phenomenology. Not only has the concept of consciousness been a core element for the philosophy of Edmund Husserl (the founder of phenomenology), but the same concept has also been developed in various directions, including one that links the consciousness issue with our body issue.

A Matter of Carnality

We would like to spend a few words about French philosopher Maurice Merleau-Ponty, because this will take us directly into our main topic of interest here: the *body issue*. We just argued in favour of the richness of a direct link with the world, as allowed by consciousness. Merleau-Ponty,[14] as a phenomenologist, agrees with the idea that such a connection and its qualitative tissue is paramount for the way in which we understand our world. Furthermore, he emphasises the importance of our body in this process. Indeed, we get to know our world via our senses, which are incarnated in our body, made in a certain way and capable of giving us the very peculiarly human quality of existence. We do not know our world as a sequence of 0s and 1s, yet in a qualitative way that is very difficult (if not impossible) to formalise, and that is connected with the very way in which our senses give us an image of reality. This concerns not only the gnoseological side of our »trade« with the world, but also all that is connected with the world's semantics and so, again, the qualitative side of reality. We give meaning to our experience, and we do this via our body, via our »carnal« experience of the world.

Roland Barthes follows Merleau-Ponty's trail as he writes about *musica practica*,[15] namely our bodily experience of music, strongly connected with the physical sensations of actually playing an instrument (or singing). An intuitive example of this can be made by mentioning the concept of the »grain of the

14 See e.g. Maurice Merleau-Ponty, *Phénoménologie de la perception*, Paris : Gallimard 1945.

15 See Roland Barthes, *Image, Music, Text*, London: Fontana Press 1977, pp. 149–154.

voice«. With this, Barthes refers to the physical connotations of a vocal sound, which carries a certain meaning that we can understand only because we know how it feels to produce such sounds, physically and corporally speaking. With the voice it is particularly evident what Barthes is talking about, since everyone knows, at least a little bit, what it feels like to sing, and so we can easily get the physical connotation of such an act. We can feel the vibration of vocal cords and the tension of muscles implied in every sound produced by a singer (this is especially clear if we think of distorted styles of singing which can hit us with particular strength). Needless to say, this is impossible for a computer, which does not have a body. And even if it had a body, it would only be able to collect perceptions later to be translated into 0s and 1s devoid of meaning and of qualitative traits.

If these claims sound too mystical – after all, is it not true that our nervous system works with 0s and 1s as well, in its own way? – maybe it will be interesting to confront them with some of the latest discoveries in the field of neuroscience. Indeed, they have a lot to share with the concepts of embodied cognition and embodied simulation,[16] employing mirror neurons to explain why we can »connect« in a special way with the goal-oriented actions we watch or listen to. In fact, mirror neurons are neurons activating when we witness someone else doing actions that would require those same neural resources, were we the ones doing them. Thus, we are able to somewhat »feel« like we were doing things other people are doing. This (relatively) recent discovery by Vittorio Gallese (2005), based on previous research made by Giacomo Rizzolatti and his team of neuroscientists during the 1990s, when mirror neurons were discovered in the first place, has fostered the development of new approaches for studying empathy, the concepts of embodied cognition and simulation being some of the most interesting results. Originally elaborated by Vittorio Gallese, embodied simulation assumes that mirror neurons can be used to »simulate« within ourselves what other people feel, and how it feels to do what they are doing. Curiously, because of the way mirror neurons work, this is only possible when actions are recognised as intentional and teleologically oriented actions. Mirror neurons can identify the goal behind actions, to the point of not activating when witnessing action B, if the performed movement is the same as in

16 See Vittorio Gallese, »Embodied Simulation: From Neurons to Phenomenal Experience«, in: *Phenomenology and the Cognitive Sciences* 4/1 (2005), pp. 23–48; Vittorio Gallese and Corrado Sinigaglia, »What Is So Special About Embodied Simulation?«, in: *Trends in Cognitive Sciences* 15/11 (2011), pp. 512–519.

action A, but the goal is different.[17] This implies that a certain »vocabulary of acts«, also containing teleological information on such acts, must be known in order for us to resonate with other people's actions in a proper way, and this anticipates an »open issue« to which we will return later.

According to Katie Overy and Istvan Molnar-Szakacs, since music »involves the perception of purposeful, intentional and organised sequences of motor acts as the cause of temporally synchronous auditory information«,[18] our experience of music can trigger that particular corporeal simulation allowed by mirror neurons, making us feel, to some extent, sensations that are related with the production of the music we are listening to and its gestures, in all their quality and meaning – almost as if we were the ones producing that same sound. This emphasises the importance of concretely »dealing with music«,[19] focusing on our living bodies and their relationship with musical instruments as interfaces between ourselves and the world of sounds.

Open Issues

The way in which embodied cognition seems to be related with music opens up several topics of discussion, the most urgent being, perhaps, what we call the »expertise issue«. Indeed, in order to physically understand music, especially when it comes to the instrumental side of it, a sort of repertoire of acts[20] – of physical gestures – must apparently be known by the listener for everything to work in the right way, as we suggested above. This means that if I do not know how to play guitar, I will never get a full experience of »how it feels« to play a specific guitar solo, not only because I have never felt that sensation in my hands, but also because it is not guaranteed that I know *what* kind of movements are required for the performance of a specific lick. If I am not an expert

17 See Giacomo Rizzolatti and Corrado Sinigaglia, *So quel che fai: Il cervello che agisce e i neuroni specchio*, Milano: Cortina 2006, pp. 23–51.

18 Istvan Molnar-Szakacs and Katie Overy, »Music and Mirror Neurons: From Motion to ›E‹motion«, in: *Social Cognitive and Affective Neuroscience* 1/3 (2006), pp. 235–241, here p. 236.

19 Mark Reybrouck, »Music Cognition and the Bodily Approach: Musical Instruments as Tools for Musical Semantics«, in: *Contemporary Music Review* 25/1–2 (2006), pp. 59–68, here p. 62.

20 See Marc Leman, *Embodied Music Cognition and Mediation Technology*, Cambridge: MIT Press 2007.

listener, moreover, I may not be able to physically resonate with a lot of elements, not only because I have never played any of those instruments, but also because I might not be capable of recognising the instruments that are being played from their sounds. For this reason, Molnar-Szakacs and Overy[21] argue that there is a sort of hierarchy of physical understanding of music, involving: 1. intention level; 2. goal level; 3. kinematic level; 4. muscle level. This explains why people resonate to different degrees with music: only musicians can resonate with it up to the muscle level,[22] but the casual listener can still understand something. For example they could feel the beat, follow the voice, perceive the connotations of very basic parameters (like pitch height, speed or intensity) and do some »mimetic subvocalisation«,[23] that is the act of vocally reproducing the music, that already creates some physical connection with what happens in the composition. This also implies that an AI cannot resonate with music at any level. Making music does not only employ calculation or specific knowledge, but most of all involves performativity, especially when it comes to popular music, that is usually conceived in the very act of playing the instrument.[24] This also means that compositional choices are often taken from the initial affordances of the instrument, and that specific techniques imply specific movements that can carry additional layers of meaning.

How much of the meaning and strength of a typically post-rock section performed in tremolo picking is carried by the resonance with such a physically impactful movement? Consider a track like »Nostalgia« by the Japanese post-rock band Mono, from the album *For My Parents*.[25] Most of the track is based on a sequence of melodic lines performed by an increasingly disruptive guitar played in tremolo picking. Especially in the first minutes, in which the guitar sounds are still quite clean, you can clearly hear the sound of the pick every single time it hits the strings, producing the distinctive »continuous« sound of the tremolo picking technique. Can only guitarists resonate with what the player is doing here? Are they the only ones who can feel the fatigue in their

21 See Istvan Molnar-Szakacs and Katie Overy, »Being Together in Time: Musical Experience and the Mirror Neuron System«, in: *Social Cognitive and Affective Neuroscience* 26/5 (2009), pp. 489–504.

22 See Leman, *Embodied Music Cognition and Mediation Technology* (note 20).

23 See Arnie Cox, *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*, Bloomington: Indiana University Press 2016.

24 See Allan F. Moore, *Rock, the Primary Text: Developing a Musicology of Rock*, Hants: Ashgate ²2001, pp. 59 f.

25 Mono, *For My Parents*, Temporary Residence Limited TRR213CD, 2012.

picking hand, as they know far too well how exhausting it can be to play in tremolo picking for such a long time? And how much of the meaning and emotional impact of this track is conveyed by this specific motor element? Current research on embodied simulation in music would probably confirm the emotional strength of the recorded action, but it would also be sceptical about the possibility average listeners have of resonating up to the muscle level with it. After all, guessing which technique is involved in playing a specific part is not an obvious task if I do not play that instrument. And it is even less obvious to know »how it feels«, physically, to play such a part.

Another possible problem exists: if that kind of resonance is strictly connected with the emotional feedback I get from music and with a level of comprehension in which corporeality and semantics intertwine, how can I get any emotional feedback from music that is not actually played by anyone? This problem concerns a lot of electronic music, and in fact many scholars acknowledge this problem,[26] but it is easy to think of musical examples of electronic music that are experienced as very physically connotated, although what we are hearing was probably programmed instead of performed in a traditional sense. What about a track like »Emerald Rush« by Jon Hopkins, from the album *Singularity*?[27] It is hard to say that this kind of music is cold, abstract, not physically nor emotively connotated in any way. This is especially true in a track like the one we mentioned, in which the rhythmic element is very present, and the sound fabric adds a kind of »organicity« to the overall result. Moreover, more and more music is now recorded using virtual studio technologies (VSTs) that are ultimately made of samples recorded separately, and it can sound »as if« played by someone while being totally assembled in the studio by programming the sequencer in an appropriate way. If we admit that our way of feeling music and »resonating« with it emotionally is bound to the motor acts behind its creation, we must try to explain how disembodied music can touch us. According to Arnie Cox,[28] there is a hierarchy of sounds to which we can physically participate with more or less ease. Electronic sounds

26 See e.g. Greg Corness, »The Musical Experience through the Lens of Embodiment«, in: *Leonardo Music Journal* 18 (2008), pp. 21–24; Cox, *Music and Embodied Cognition* (note 23); Leman, *Embodied Music Cognition and Mediation Technology* (note 20); Molnar-Szakacs / Overy, »Being Together in Time« (note 21); Reybrouck, »Music Cognition and the Bodily Approach« (note 19).

27 Jon Hopkins, *Singularity*, Domino WIGCD352, 2018.

28 See Cox, *Music and Embodied Cognition* (note 23).

can be found in the middle of the hierarchy, and actually two variants are there as well: electronic sounds produced via hand controllers or electronic sounds generated and modified via real-time controllers (like a DJ would do). Once again, the issue seems to find a possible solution in the implementation of a gradation in the strength of the experience. Though being less physical, electronic sounds may be still possible to embody, to some extent.

However, these examples had to be mentioned just to emphasise the fact that there is still much to be understood in this field. They do not, in our opinion, affect the validity of our main claim about the body issue: a machine misses all the layers of meaning and feeling connected with physicality, with our understanding of things from the perspective of embodied conscious beings. Let us forget our anti-Terminator weapons because it looks like we overestimated our competitor's possibilities. At the very least, this misplaced urge can suggest to us innovative ways for understanding things as complex as the way in which we resonate with the physical side of music, which, for some, might not be as fascinating as science fiction, yet we think it is just as mysterious and amazing, and much more real and relevant.