Eduardo Reck Miranda *Editor*

Handbook of Artificial Intelligence for Music

Foundations, Advanced Approaches, and Developments for Creativity

Springer

Handbook of Artificial Intelligence for Music

Eduardo Reck Miranda
Editor

Handbook of Artificial Intelligence for Music

Foundations, Advanced Approaches, and Developments for Creativity

Springer

Editor
Eduardo Reck Miranda
Interdisciplinary Centre for Computer Music
Research (ICCMR)
University of Plymouth
Plymouth, UK

ISBN 978-3-030-72115-2   ISBN 978-3-030-72116-9 (eBook)
https://doi.org/10.1007/978-3-030-72116-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword: From Audio Signals to Musical Meaning

In 1957, Lejaren Hiller and Leonard Isaacson stunned the world of music by presenting the first composition constructed by an AI system, called *Illiac Suite* (Hiller and Isaacson 1959).

Illiac was the name of one of the first computers ever built, installed at the University of Illinois in 1952. Hiller and Isaacson were trained as musicians but they were also computer scientists 'avant la lettre' with a solid training in the natural sciences. Their project took place in the wake of the earliest problem-solving programs demonstrated by Allen Newell, Herbert Simon and John Shaw a few years earlier and the enthusiasm generated by John McCarthy and Marvin Minsky at the Dartmouth summer project on AI in 1956 (Nilson, 2010).

Illiac Suite is remarkable from many angles, particularly given the state of computer and software technology at that time. Programs had to be submitted on punched cards, memory was tiny, execution slow, and higher-level programming languages were in their infancy; the first compiler for Fortran became operational only in 1957. One had to be a genius to get anything done at all. The *Illiac Suite* composition was also remarkable because Hiller and Isaacson introduced various paradigms for computer music that are still dominant today.

They were familiar with the canonical techniques of Western composition based on a system of pitches and constraints on how these pitches could best be organized to get a harmonious piece of music, such as use recognizable tonalities, avoid transitions that are boring like parallel fifths or octaves, and so on. Within the heuristic search paradigm initiated by Allen Newell and Herbert Simon (Newell and Simon, 1956), Hiller and Isaacson implemented a generate and test scenario where possible pitches were generated and only those kept that fit with the canonical rules of composition, thus foreshadowing the constraint-based computer composition techniques used today.

To dampen the inevitable combinatorial explosions, they introduced heuristics and higher-level musical representations, such as larger melodic and rhythmic structures. They also experimented with Markovian decision processes that are still at the heart of many efforts in computer music generation and argued that creativity could be modelled by introducing stochasticity, based on a random number generator that could make unexpected choices. To bridge the gap with human-produced music, they introduced higher level structures from the classical music tradition, for example, having different movements like presto, andante, and allegro, and they used human performers and classical instruments to give a recognizable emotional quality to their music.

How did the resulting music sound? You can listen for yourself.[1] The music is certainly intriguing and an adequate performance can add emotional value making it easier (or in some cases harder) for listeners to build some sort of interpretive experience. After all, enjoyable music is not only the task of the composer and the performers but just as much of the listeners who are invited to project structures and meanings on what they are hearing.

At the time, most musicians reacted in an extremely hostile way to this experiment, both because the computer was encroaching on a terrain that was until then the province of human creativity and because of the aesthetics and structure of the music that the computer programs produced. With respect to the latter, we have to remember that the *IlliacSuite* was composed at the time John Cage's experimental music had come in vogue, emphasizing aleatoric elements, processes, and rhythm and tone rather than melody and harmony (Kuhn 2016). From that perspective, *Illiac Suite* is actually more conservative than much of the experimental music, concrete and electronic music that followed. It is more comprehensible to the average listener than the highly complex academic music produced in the 1980s and 90s. Nevertheless, *Illiac Suite* remained an isolated experiment.

Fast forward to today. The field of computer music in general, and the application of AI in music in particular, have blossomed beyond belief. Developments at the level of hardware, software, and the use of AI for composition, tutoring, recording and music distribution have been extraordinary. The papers appearing in the *Computer Music Journal* (started in 1977) and the historical collection of Stephan Schwanauer and David Levitt (1993) are important resources to track these developments. And now we have the collection of chapters in this handbook, brilliantly brought together by Eduardo Reck Miranda. This book gives an excellent survey of more recent achievements and speculates on near-future developments. All the publications, demonstrations and musical works being discussed here establish beyond doubt the very high level of technical and scientific competence and the musicological sophistication of the computer music field and its branch dedicated to the application of AI to music.

The experimental achievements of computer music researchers have been abundant. But equally impressive is the fact that the laboratory experiments have successfully moved into musical practice in the real world. No composer today would work without the help of programs for editing scores and for the tedious process of adapting scores to different instruments. Synthesis from scores has become so good (but certainly not at the level of human performers) to give composers a good idea of what their music will sound when executed by a human orchestra. Performers now practice and play from digital scores and have digital ways to organise and annotate scores. They can even practice on their own with the

[1]https://www.youtube.com/watch?v=n0njBFLQSk8 (Accessed on 02 February 2021).

other instruments synthesized by a computer, in real-time synchrony with what they are playing. The recording and distribution of music are no longer the exclusive work of manual labour only because very sophisticated signal processing intervenes in recording and listening. Moreover, modern-day AI algorithms play a crucial role in how the public gets to know new work. Also in musical education, the results from computer music research are playing increasing roles from simple Apps that help you train your musical competence to MOOCs that allow many more people to learn about music online (Steels, 2015). And moreover, we have many examples of fascinating music composed, and in many cases performed, by computer systems in prestigious venues and launched on a commercial market where until recently only human compositions could be found. This is all very remarkable. The adoption of computer music has even accelerated with the COVID-19 pandemic in 2020, so that you now find choirs that rehearse using Internet streaming technologies, or Jazz musicians that play together over the cloud, with occasionally an AI musician thrown in to play along when a band member is missing.

Research into computer music has not only contributed to music itself. Computer music researchers have also been making steady contributions to software engineering, hardware development, signal processing, embedded systems, and AI methods, particularly in the area of constraint programming, object-oriented programming and, more recently, deep learning. At the moment there are even forays into neuro-technology (integration of electronics and computing with living neural cells) and quantum computing (see Chaps. 8 and 34 this volume). All of these topics are explored in this book with outstanding overviews and reviews.

Given all these incredible advances what could be done next? Pushing the state of the art further in computational creativity is high on the list, and one of the focal topics of the present book, so let me focus on that for the remainder of this essay.

So far most experiments in musical creativity still follow the approach of Hiller and Isaacson, namely to work with templates and add some randomness in instantiating a template to make an audible composition. The difference nowadays is that those templates might be learned using machine learning techniques and the randomness gets heavily constrained both by statistical models and by music theory so that the resulting music often sounds more harmonious and plausible than *Illiac Suite*. But is this the only approach possible?

Although composers certainly use templates that they have either learned explicitly through musical education or implicitly by listening to a lot of music, their creativity obviously goes beyond making random variations to instances of templates or to the templates themselves. And although listeners also have musical memories containing templates, melodies, favourite interpretations of well-known pieces, and so on, they clearly do more than recognize patterns and predict what pattern comes next. I am of the opinion that in order to know what it is that they do more, and hence what creativity requires in the domain of music, we have to address the question of meaning in music. I am probably stating the obvious when I say that *musicians and listeners engage with music because they find music meaningful*.

What are the implications of saying that? Does it make sense actually to talk about meaning in music? What kinds of meanings are we talking about? And how does music accomplish the expression of meaning? How do listeners ‘understand’ music? Can we build tools to assist composers for the meaning dimension of their work? Can we find computational means to aid listeners in interpreting and experiencing in a richer and deeper way the meanings of a musical work? Most importantly, if we are serious about music creation or co-creation with AI systems, should we address how musicians use music to express meaning? And furthermore, if we are serious about autonomous computational creativity for music, should we investigate how innovative composers or improvisers have found new ways to express meaning through music or opened up new domains of meaning for expression by music?

All these questions are difficult, which might explain why the AI and music community has been avoiding them. In fact, recent work on machine learning and neural networks has moved AI even further away from considerations of meaning because of its behaviourist approach to mind. In the behaviourist tradition, initiated by Watson (1930) and Skinner (1953) and brought into AI in the 1950s by Frank Rosenblatt (1962) and his cybernetic colleagues, intelligence is reduced to pattern recognition and pattern prediction It is acquired through behavioural conditioning, associative learning, reinforcement learning, perceptron-style multi-layered networks and other statistical machine learning techniques. Goals, intentions, beliefs, symbols and perspectives are de-emphasized, and occasionally denied to be relevant for intelligence. The topic of meaning is avoided entirely.

We see this kind of approach not only in applications of AI to music but also for all other applications areas that this kind of data-driven, behaviourist AI has tackled. For example, ‘neural’ text translation approaches translation by mapping n-grams (sequences of words) from a source to a target language without even trying to do a serious syntactic and semantic analysis, let alone try to understand and reformulate what the author wants to say. Adepts of this approach argue that doing syntactic and semantic analysis is very difficult, which is certainly true, but also that you do not need it. When more data is given to learn from, the glitches and silly translations that we now see will become rare. Another example is ‘neural’ art generation (Dumoulin et al. 2017). A neural network learning system generates paintings in the style of the impressionist Claude Monet, for example, by regenerating statistical patterns gleaned from Monet’s paintings, but without even trying to understand what is depicted on the painting. The system would not have a clue as to why a particular scene was chosen or why Claude Monet made a specific transformation of the original colours and shapes.

We see this same approach in data-driven AI work that explores how deep learning mechanisms can be used for music, for example, in the experiments of MUSENET by OpenAI.[2] Even though it is quite fascinating and a technical tour de force, attempts to create music ‘in the same style’ as human composers like Chopin,

[2]https://openai.com/blog/musenet/ (Accessed on 02 February 2021).

based on a probabilistic transformer model that predicts or generates the next element in a sequence, completely bypasses meaning. What is amazing is that the end result is so close to a Chopin piano piece that commentators often describe it as 'brilliant', but also 'terrifying' and 'having no human element in it'. It suggests that a significant part of the musical experience is imposed by listeners. The composer creates a vehicle for this to happen. And even if the musical piece has been constructed without any consideration of meaning, listeners still manage to impose structure and meaning onto it.

In contrast to the behaviourist tradition, the cognitivist tradition, brought into AI by Allen Newell and Herbert Simon (1956), insisted on topics like goal-directed problem solving, symbolic representations, reasoning, conceptual understanding and Piagetian style constructivist learning. Some research on understanding did indeed take place in the 1960s and 70s, see, for example, Minsky (1968) or Schank (1983), but in general, the cognitivist AI tradition has also tended to shy away from meaning. They have focused instead on syntactic manipulation of formal structures, simply because this is more amenable to computational treatment. The composition and analysis experiments to formalise and codify music using symbolic techniques (see, for example, Chap. 18 in this volume) are technically as impressive as the more recent music generation experiments based on machine learning, particularly when used for real-time support during Jazz improvisation. The results also sound more like real music, even though discerning listeners might still feel a lack of narrative structure, authentic emotions and meaning. Nevertheless, the cognitive AI tradition accepts at least that meaning is important in human intelligence and culture.

Computational music research on meaning has been difficult because there is no clear consensus among musical scholars and practitioners on what musical meaning is. Even the idea that music is about meaning is controversial. For some, there is simply no meaning in music. For instance, consider this quotation from composer Igor Stravinsky: "I consider that music is, by its very nature, essentially powerless to express anything at all, whether a feeling, an attitude of mind, a psychological mood, a phenomenon of nature, etc." (Stravinsky, 1935, p. 53). This quotation is surprising, particularly coming from a composer that has written music for ballet, a few operas, and even music for film. If we take this stance, music becomes similar to mathematics. Musical composition comes to be seen as about designing abstract structures and formal pattern manipulation. Listening in this case amounts to an experience of sound and the recognition of the patterns imposed on sounds. From this perspective, the joy of musical listening consists in recognizing and tracing the development of these structural patterns, similar to looking at mosaic patterns on a floor.

The 'music is like mathematics' metaphor is natural to the many mathematicians and computer scientists, who have been the most active group in computer music research. They feel very much at home in the world of abstract structures and the computer is the ideal tool for exploring this world. This perspective has heavily influenced the development of AI as applied to music. However, equating music with mathematics is nor the view of many practicing musicians nor of most listeners. For them, music is much more than sound sensations and syntactic structures. We are a meaning-seeking species steeped in emotion and intention. We are always trying to figure out why things are the way they are, what motivates somebody to do something, how fragments of an experience fit together into a coherent whole and relate to our earlier experiences.

True, many of the meanings that we impose on reality and on artworks, including music, are not easy to capture in words. They are pre-verbal and non-symbolic, but they still count as meanings. It is also true that even if we would have a clear notion of what musical meaning might be, the set of meanings invoked by one listener would seldom be the same as those invoked by another listener. This is simply because different people have their own personal memory, their own prior experiences of the world and of music, their own social context and psychological state when composing or listening to music. Objectively, there is no 'correct' set of meanings for a given piece of music. Therefore, it does not seem to be a reasonable goal for AI to extract it. It also seems beyond machines to capture the rich embodied and culturally grounded set of meanings that humans effortlessly deal with.

Those who admit that music is meaningful often restrict the meaning of music to be about expressing and invoking emotional states, like sadness or joy (Meyer, 1956). That is certainly one aspect of musical meaning, but there is much more. I suggest that it is helpful to look at other artistic disciplines that have been grappling with meaning in art.

One concept that I have found useful is that of a narrative. It is commonly employed in studies of art and literature. A narrative is a larger scale structure that organizes experiences into multiple levels of description. A painting or a musical composition is not literally a story, like a theatre play or a novel, but it stimulates us to construct narratives. From this perspective, an artist is engaged in a form of cognitive engineering (Dewey, 2018), manipulating the mental processes of viewers by shaping their sensory experiences and memory recalls in order to stimulate narrative construction. This insight is very important because it suggests that a composer or performer is like a designer who has goals at many levels and almost magically manages to transform these goals into a coherent piece of art. It also suggests a rethink of how we might achieve computational creativity in the music domain. It is about finding solutions to compositional problems similar to the way an architect designs a building that has to satisfy many constraints and at the same time has to be done in a creative way. Creativity then is not about introducing some random variations without motivation or insight into the underlying purposes and strategies for achieving them.

Narratives typically segment experiences into a series of events with a temporal, causal, and spatial structure. They describe the different participants in these events their roles, goals, intentions and emotions. They introduce the context and world setting, provide perspectives, an emotional stance, and probe the moral, political and ethical implications of what is happening. When we look intently at a painting or watch a theatre play, we spontaneously construct narratives. We try to fit together the different elements we see or hear until they fall into place. An artistic experience, or even a mundane situation, only 'makes sense' when we can construct a coherent narrative and integrate it with our own personal episodic and semantic memory. Often there are ambiguities, alternative competing interpretations and conflicting perspectives which either get resolved or remain as open-ended multiplicities. As composer Arnold Schoenberg puts it: "A work of art is a labyrinth where, at every point, every knowledgeable person is aware of the entrance and the exit without needing a golden thread to guide him." (Schoenberg, 1995). I think Schoenberg overestimates most listeners who may struggle to hear the structure underlying his 12-tone compositions, but the point is well taken.

Another, complementary, insight which I have found useful, comes from the art historian and semiotician Ervin Panofsky. He has identified five levels of meaning (Panofsky, 1939): the formal, factual, expressional, cultural, and intrinsic level, and applied it principally to painting. Can we apply these Panofskyan distinctions to music as a step towards putting meaning at the heart of AI research into music? Let me try and see what lessons we can learn.

The first Panofskyan level is that of the *form*, the material presence and syntactic structure of an artwork. For a painting, these are the lines and colours which hierarchically aggregate into artistic motifs. The obvious correlation for music are the sounds themselves, which have sensory qualities that may already give aesthetic sensations or a feeling of well-being, and the syntactic structuring of the sounds: their segmentation, categorization and aggregation into tones, melodies, rhythm, meter, harmonic structure, phrases, and the like. They constitute musical motifs, or musical ideas in Schoenberg's terminology (Schoenberg, 1995). They are the building blocks of a composition at the form level.

Composers often already tell a story on this form level, playing around with these musical motifs. They are presented, transformed, repeated and contrasted with other motifs as the music unfolds. Music is unique as an art medium because the musical forms themselves create narratives that are about musical ideas without any reference to emotions or events in the world. Minsky's brilliant essay *Music, mind and meaning* (Minsky, 1981) illustrates this point, using the example of Ludwig van Beethoven's fifth symphony, where the first subject is expressed in its famous first four notes; see also (Guerrieri, 2012). If some musicians, such as Stravinsky, say that music does not have any meaning, I reckon that what they want to convey is that there is no meaning outside of the domain of (musical) form itself in music. I am not entirely convinced about this.

The level of musical form is what most of the AI research into musical composition, interpretation and listening has focused on. Much has been achieved, as the chapters in this book clearly attest, and there are still many ideas floating around on how one can increase the structural depth of compositional work, build better ways to extract from audio signals notes, tempo, measures, rhythms, and harmonic structure or improve the synthesis of music by taking into account its phrasal structure. However, less work has been done on the narrative structure of music at this form level: to recognize the motifs and their transformations, to reconstruct the musical story that the composer is telling us.

The second Panofskyan level is concerned with *factual meanings*. It refers to the capability of images to conjure up memories of objects and events in a real or imaginary world. For example, Michelangelo Merisi da Caravaggio's stunning painting *Presa di Cristo nell'orto* (*The Taking of Christ*, in English) from 1602 directly triggers recognition of a central narrative in the catholic passion story that would have been known by his audience at the time.[3] In the painting we see in the middle Judas giving the fatal kiss and the soldiers ready to take Jesus. To the left, we see one of the apostles crying for help and Jesus retracting to avoid the embrace of Judas. To the far right, we see a person (generally considered to be the painter himself) who shines light on the scene, thus becoming an accomplice to the arrest. The title of a work is usually suggestive of its factual meaning. But it is only the starting point, together with the image itself, for triggering the construction of a narrative that makes sense of what is depicted and why.

Factual meanings are much more common in music than usually thought, although they are less so in twentieth-century music, which commensurate with the rise of abstract art by painters such as Piet Mondriaan and Wassily Kandinsky who wanted to create paintings who were interesting at the level of form only, using music as their guiding example. A figurative component is most evident for vocal music, where the music underscores and augments the verbally expressed narrative, or for opera and film music, where the music underlines the action, the emotional state of the characters and their role.

For example, Johann Sebastian Bach's *St. Matthew Passion* composed in 1727 tells the same story as Carravagio's painting and is equally figurative. The story of the arrest of Christ is first told by the Evangelist without much drama (Recitativo 32). Initially, there is a duet (Duetto 33) with a feeling of sadness and resignation with the choir representing the apostles and interjecting with the cries "Laszt ihn, haltet, bindet nicht!" ("Leave him, don't' keep him, do not tie him!").[4] These cries map straight onto the left-most figure in Caravaggio's painting. But then a storm breaks out: "Sind Blitze und Donner in Wolken verschwunden" ("Lightning and thunder disappear in the clouds"). It is forcefully evoked by the choir in staccato rhythm and totally dramatizes the importance of the arrest.[5] In order to appreciate all this, one has to go beyond the form appearance of this music and take into account what the story is about.

Factual meanings are not only present in vocal works. Music does not necessarily have to imitate literally the sound of an event or situation in the world to be figurative. It is most of the time only suggestive, the same way a painting of the sea by impressionist painter Claude Monnet does not literally reuse the colours of the sea and the sky, or faithfully represents the waves that you actually observe. The

[3] https://en.wikipedia.org/wiki/The_Taking_of_Christ_(Caravaggio)#/media/File:The_Taking_of_Christ-Caravaggio_(c.1602).jpg (Accessed on 04 February 2021).

[4] Phlippe Herreweghe, Collegium Vocale: https://www.youtube.com/watch?v=70shLtLxcYA (Accessed on 04 February 2021) Storm starts at 3:33.

[5] Jos van Veldhoven, Nederlandse Bachvereniging: https://www.youtube.com/watch?v=_uLpp6cW7sA (Accessed on 04 February 2021).

relation to reality is iconic: there is a resemblance with what is signified without trying to be realistic. Take Antonio Vivaldi's *Four Seasons* violin concerti composed in 1723. You can of course listen to it purely from a form point of view and discern different musical motifs, and hear how they develop and interact over time. But this ignores that the work is really about the four seasons: Summer, Spring, Winter, and Autumn. Each season is evoked with musical images conjuring up impressions and experiences related to that season. For example, in the Summer movement[6] we hear the laziness that comes with a sizzling hot sun, water flowing, birds, a barking dog, buzzing flies, but also the wind coming up, lightning bolts, thunder and then a violent escalating storm. Is this over-interpretation? Not really. Vivaldi himself wrote sonnets corresponding to each movement. The sonnets describe what experiences he was trying to evoke, helping the listener's imagination. Without considering this figurative aspect, listening to *Four Seasons* becomes an exercise in syntactic recognition that lacks meaning and therefore becomes boring once these sounds and structures have been grasped at their surface level.

Other examples of factual meaning in figurative music are easy to find. Claude Debussy's symphonic sketch *La Mer* (*The Sea*, in English) composed in 1903 is really about the sea. The first movement *De l'aube a midi sur la mer* (*From dawn to midday at the sea*, in English) gives the sensation of the swaying movement of the waves that start timidly but then become bigger with the water splashing as they break.[7] Another example are the so-called tone poems by Richard Strauss. For example, the last one of the *Vier letzte lieder* (*The four last songs*, in English) called *Im Abendrot* (*At Sunset* or more literally *'With the red of the evening'*, in English) was written in 1948 shortly before the composer died.[8] The song for soprano and orchestra evokes the red light of the sun going under but also the end of life depicted as a long walk by a couple. When the soprano sings 'Zwei Lerchen nur noch steigen/nachträumend in den Duft' ('Left are only two larks who climb like in a night dream in the air', in English), we hear singing larks evoked through two flutes. These larks are metaphors for the souls of the couple that are soon to go to heaven. The last phrase is: 'Wie sind wir wandermüde-Ist dies etwa der Tod?' ('We are tired of walking-is this near death?', in English) where Strauss reuses the basic theme of 'Tod und Erklaerung' ('Death and transfiguration', in English), another tone poem he wrote 60 years earlier, which is a musical portrait of a dying man entering into heaven. Clearly, when this broader context is provided the experience of *Im Abendrot* changes completely and goes beyond its remarkable sensual beauty.

The third Panofskyan level is about *expressional meaning*, in other words the psychological states, emotions and affects that are evoked through an artwork. Looking again at Caravaggio's *Presa di Cristo nell'orto* we see that every figure

[6]Mari Samuelsen, Trondheim solists: https://www.youtube.com/watch?v=g65oWFMSoK0 (Accesses on 04 February 2021).

[7]Claudio Abbado, Luzern Festival Orchestra: https://www.youtube.com/watch?v=SgSNgzA37To (Accessed on 04 February 2021).

[8]Anja Harteros, Sinfonieorchester des Bayerischen Rundfunks: https://www.youtube.com/watch?v=JwZOXC6_4fE (Accessed on 04 February 2021).

expresses clearly emotional states through the looks of their face, the gestures and body language. The apostle expressing fear while crying for help, Jesus almost saying to Judas: ‘What are you doing?’, frowning and looking sad, the onlooker to the right being curious and astonished about what is going on. Obviously, expressional meanings are abundant in music as well, and I am not referring now to expressive performance that brings out the music’s phrase structure, which is on the form level, but to the expression of affective states, often related to the factual meanings of the musical work. Such expressive meanings are embedded in all the aspects of a piece of music: the tonality, the chords and chord progressions, the choice of instruments, the tempo, dynamics, loudness and articulation. The responsibility for recognizing the affective content in a score and expressing it in performance lies for a large part in the hands of human interpreters. It is the Achilles’ heel of synthetic music.

The song by Richard Strauss showed already a magnificent example of a strongly emotional content. But let us listen to George Friedric Haendel’s famous aria *Lascia ch’io pianga* from his opera *Rinaldo*[9] as a second example. The aria, composed in 1685 is sung by the character Almerina, who is held captive, away from her lover Rinaldo. The words are as follows: ‘Lascia ch’io pianga / Mia cruda sorte, / E che sospiri / La libert’a Il duolo infranga / Queste ritorte, / De’ miei martiri / Sol per pietà’ (which translates in English as ‘Allow that I weep over / my cruel fate, / and that I may sigh / for freedom. Let my sadness shatter / these chains / of my suffering, / if only out of pity’). The aria, in a tonality of F-major, and its orchestration is of utmost simplicity and you do not need to follow or understand the words to share the extreme feeling of sadness of Almerina, the desperation and frustration of being away from her lover, the longing for liberty. Many other magnificent examples of the importance and power of expression in music abound, including in purely instrumental music, and to ignore it is to deprive music of its potent force.

The expressive meaning level has already received considerable attention within AI approaches to musical meaning; see, for example, Widmer and Goebl (2004) and Chap. 19 this volume, because it is so crucial for a good performance. But so far, expression is mostly considered from the viewpoint of making the syntactic structure of the music more legible. This is very important in itself of course, but the expression of affective states and how they embed in a narrative remains almost virgin territory for AI approaches to music.

Next, there is the *cultural meaning* layer. It rests on knowing more about historical events, society, religious systems, myths, other cultural artefacts. For example, the Caravaggio *Presa di Cristo nell’orto* or Bach’s *Matthäus Passion* can only be understood when knowing the Bible’s narrative of these depicted events. For Christian believers, it has even more significance. They will feel total empathy with the suffering of Christ, the shock of the arrest, the protests of the apostles. Or

[9] Joyce Di Donato, Maxim Emelyanychev, *Il Pomo d’Oro*: https://www.youtube.com/watch?v=PrJTmpt43hg (Accessed on 04 February 2021).

consider, for example, the song *To Yelasto Pedi* (*The jovial boy*, in English) composed by Mikis Theodorakis, and performed to a full stadium after the fall of the military dictatorship in Greece in 1974.[10] This song, about the fascist terror exerted on the Greek population, was of enormous significance for the audience, particularly because they were forbidden and Theodorakis was jailed and forced into exile. The unforgettable concert in 1974 was a celebration of the regained freedom.

Similarly, fully experiencing the aria *Va pensiero* from Giuseppe Verdi's opera *Nabucco*, which has become something like an alternative national hymn of Italy, can only be done by understanding that it was written when Lombardy and Veneto were occupied by the Habsburg empire in the nineteenth century and the population felt enslaved like the slaves depicted in the opera. When this aria is performed today, Italian audiences often rise to sing along. The aria resonates because it was at the time a symbol of protest against oppression by the Habsburg armies. More recently it has been appropriated by Italian right-wing parties because of its patriotic symbolism but Ricardo Mutti in 2011 invited the audience to re-appropriate this highly symbolic music to protest against the dismantlement of cultural institutions and to regroup the cultural force of Italy and thus halt the slide in moral decay of Italy during the era of Silvio Berlusconi.[11]

Here is another example: the Jazz standard *Strange Fruit*. You can certainly listen to it—and it is often performed that way—as just another Jazz standard where the different performers give their own interpretations of the melody and harmony. However, this is totally missing the cultural and political significance of this song. It is actually a cry against the injustice of racial oppression that was happening in the south of the US in the 1920s and 30s when this song was conceived and performed. *Strange Fruit* refers to bodies of black people that have been tortured and lynched: 'Southern trees bear a strange fruit / Blood on the leaves and blood at the root / Black bodies swingin' in the Southern breeze / Strange fruit hangin' from the poplar trees'. When watching Billy Holliday's rendition of this song[12] there is raw authentic emotion. At the time, it raised high tensions and there were attempts to silence her, particularly as she sang it before white audiences. Without this historical context, the significance of the song is largely lost.

It is clearly too much to expect from a computer program to take this cultural context into account or to construct an authentic performance with the same cultural background and emotional force that humans bring to bear on such musical experiences. Nevertheless, this is what music in its full extent is all about. As more and more cultural knowledge becomes available through the World Wide Web and as the knowledge graphs and dictionaries (e.g., Wordnet or Propbank) have now

[10]Maria Farantouri, Mikis Theodorakis: https://www.youtube.com/watch?v=NLgerQJo7zM (Accessed on 04 February 2021).

[11]Ricardo Mutti, Scala of Milano: https://www.youtube.com/watch?v=5wAXhHrqOzQ&list=RDXg1yRoENqJQ&index=2 (access on 4 February 2021).

[12]Billy Holliday: https://www.youtube.com/watch?v=-_R8xxeMFEU (Accessed 4 February 2021).

become accessible to the AI and music research community for use in meaning-oriented AI, new opportunities will certainly emerge to preserve much more of the cultural context of musical works and to make it more accessible as part of recordings or in educational settings. That is also a task for the future.

Finally, Panofsky talks about the *intrinsic meaning* of an artwork: the motivations of the artist, what does he or she try to accomplish. This can, for example, be a political statement, social commentary, moral advice, community bonding, commemoration of traumatic events, religious and spiritual exaltation, or mere entertainment. The intrinsic meaning of Holliday's *Strange Fruit* or Theodorakis song *To Yelasto Pedi* is political, protesting against injustice. The meaning of Bach's *Matthäus Passion*, and much of his other music, is spiritual. Needless to say, this dimension is entirely lacking in computational musical compositions or synthetic performance, mainly because AI systems, despite claims in the popular press, do not have the kind of autonomous agency and social embedding that humans have. This gap is not a criticism. But it should make us all humble and critical about AI researchers claiming that their computer programs can now make music as good as human composers! They have fallen in the Turing trap: to create music superficially indistinguishable from a real composition by a (usually naïve) human observer. But this is fake music because it lacks the many levels of meaning that are the essence of human music. It is not grounded in identity or human motivation. It does not express affective values. It is not embedded in cultural and societal concerns.

Let me summarize the main point of this essay. Our current computational tools can handle a remarkable number of signal processing and feature recognition aspects for going from sounds to notes, rhythms and phrasal structures. They can represent and enact constraints on harmony, instrumentation or rhythm for musical composition. And they can even approach very difficult issues in performance and audio synthesis. Nevertheless, they rarely address the meanings and musical narratives which underlie music as an art form. Consequently, it makes no sense to call these systems creative in the same way as human composers, performers or listeners are musically creative.

I do not believe that the rich web of meanings that we as humans naturally engage in will ever be captured by an AI system, particularly if it is disembodied and has no social role in a human community. Nevertheless, the fantastic tools we have already today make it conceivable to attempt a significant leap in the direction of meaning. This development would be in line with a current trend in AI research, which considers meaning as the key barrier AI still has to overcome (Mitchell, 2020) and calls for a 'human-centric AI' where only through a proper focus on meaning can we create a more responsible and more robust form of AI than the one underlying many of today's applications (Steels, 2020).

A focus on meaning in AI research into music would allow us to understand better the relation between music and the musical narratives that it can invoke in listeners, and the strategies composers use to design music that can realize this function. I do not think this will happen soon but if it happens it would lead to many new spin-offs for enhancing musical practice, preserving or reconstructing musical heritage, richer musical compositions, and many great musical experiences for all of us. It would also allow us a whole new approach towards musical creativity, now understood as establishing multi-dimensional mappings between a complex web of meanings and musical forms.

Barcelona, Spain Luc Steels
December 2020

Acknowledgements The writing of this essay was possible by the support of the Catalan Institute for Research and Advanced Studies (ICREA) and partly funded by the European FET Proactive Project MUHAI on Meaning and Understanding in Human-centric AI. I am indebted to Oscar Vilarroya for many discussions about the role of meaning in understanding the human mind and its manifestations in art.

References

Dewey, R. (2018). *Hack the experience. New tools for artists from cognitive science*. Brainstorm books. Goleto, Ca.
Dumoulin, V., Shlens J., Kudlur M. (2017). A learned representation for artistic style. https://arxiv.org/abs/1610.07629 (Accessed on 04 February 2021)
Golio, G. (2017). *Strange Fruit. Billie Holliday and he power of a protest song*. Minneapolis: Milbrook Press.
Guerrieri, M. (2012). *The first four notes. Beethoven's fifth and the musical imagination*. New York: Random House.
Hiller, L., Isaacson L. (1959). *Experimental Music: Composition with an electronic computer*. New York: McGraw-Hill.
Kuhn, L. (Ed.) (2016). *The selected Letters of John Cage*. Middletown, CT: Wesleyan University Press.
Meyer, L. (1956). *Emotion and meaning in music*. University of Chicago Press
Minsky, M. (1968). *Semantic information processing*. Cambridge Ma: The MIT Press.
Minsky, M. (1981). Music, Mind, and Meaning. *Computer Music Journal*, 5(3).
Mitchell, M. (2020). On crashing the barrier of meaning in Artificial Intelligence. *AI magazine*, 41(2).
Newell, A., Simon, H. (1956). The logic theory machine: A complex information processing system. *IRE Transactions of Information Theory* 2:61–79.
Nilson, N. (2010). *The quest for Artificial Intelligence. A history of ideas and achievements*. Cambridge UK: Cambridge University Press.
Panofsky, E. (1939). *Studies in iconology. Humanistic themes in the art of the renaissance*. Oxford: Oxford University Press.
Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington DC: Spartan Books.
Schank, R. (1983.) *Dynamic memory a theory of reminding and learning in computers and people*. Cambridge UK: Cambridge University Press.
Schoenberg, A. (1995). *The Musical Idea and the Logic, Technique, and Art of Its Presentation. motive, gestalt, phrase, theme, rhythm, harmony, and form. English translation*. New York: Columbia University Press.
Schwanauer, S. Levitt D. (1984). *Machine models of music*. Cambridge Ma: The MIT Press.
Skinner, B. F. (1953). *Science and human behaviour*. New York: Macmillian.
Steels, L. (2015). *Music learning with Massive Open Online Courses (MOOCS)*. Amsterdam: IOS Press.

Steels, L. (2020). Personal Dynamic Memories are Necessary to Deal with Meaning and Understanding in Human-Centric AI. In: Saffiotti, A, L. Serafini and P. Lukowicz (eds). Proceedings of the First International Workshop on New Foundations for Human-Centered AI (NeHuAI) Co-located with 24th European Conference on Artificial Intelligence (ECAI 2020) CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073) Vol-2659.

Stravinsky, I. (1935). *Igor stravinsky. An autobiography*. New York: Simon and Schuster.

Watson, J. (1930). *Behaviorism*. New York, NU: Norton.

Widmer, G. Goebl W. (2004.) Computational models of expressive music performance: the state of the art. *Journal of New Music Research*. 33(3) 203–216.

Preface

I am delighted to be in a position to write this preface: it is the last task that I need to get done before I submit the manuscript for production. I have just gone through the checklist. All good to go. It was a long, but nevertheless gratifying, journey.

I must confess that I misjudged the magnitude of the job when I signed up to produce this book. It ended up being a much greater project than I had anticipated. And much harder too. The field of Artificial Intelligence (AI) today is overwhelming. It is very difficult to map. And throwing music into the mix makes things even muddier.

It was relatively straightforward to survey the field when I edited the book *Readings in Music and Artificial Intelligence* 20 years ago.[13] There were only a handful of pioneers taking music as a serious domain for AI at the time. Research into applying AI in music was in its infancy. I vividly remember the disdainful looks I used to get at international AI conferences in the 1990s when introducing myself as a musician. And to add insult to injury, musicians used to scoff at the notion of making music with AI. Perceptions have changed. And how!

Back then, neural networks were not much more than a theoretical promise based on toy problems. Practical implementations often failed to impress. Symbolic knowledge representation and logic-based modelling were the norm. Functional programming with LISP and logic programming with PROLOG defined the bastions of AI research at the time; the latter favoured in Europe, notably in France and Scotland, the former in the USA and beyond. From this era, two notable achievements immediately come to mind. David Cope's EMI system emerged as an exemplary LISP-based system, able to learn and compose in the style of classical music composers. And the Continuator system (precursor of the Flow Machines project) led by my former colleague at Sony, François Pachet, took the symbolic approach to an unprecedented level of sophistication.

Since then, neural networks evolved significantly with the emergence of the so-called 'deep learning' methods, which are remarkable. Deep learning has been enjoying considerable publicity. So much so that it has inadvertently become a synonym of AI for some: I often have students coming through the door these days

[13] https://www.taylorfrancis.com/books/readings-music-artificial-intelligence-eduardo-reck-miranda/e/10.4324/9780203059746.

thinking that AI and deep learning are the same thing. Yet, deep learning alone is of limited capacity to model and simulate intelligence; depending on what one means by 'intelligence', of course. This book concerns musical intelligence.

In addition to developing technology to building musical systems able to personify aspects of human intelligence, AI is a great tool to study musical intelligence. If anything, AI research has demonstrated that intelligence in general requires more than logical reasoning. It requires creativity, subjectivity, emotions, interaction, embodiment, and all those things that the brain takes care of, consciously and unconsciously, to keep us alive. Music engages a multitude of these human capacities. Hence the reason music has become such an interesting domain for AI research.

This book comprises 34 chapters from leading scholars and practitioners approaching AI in music from a variety of perspectives. There are chapters touching upon sociological, philosophical and musicological issues: Chapters 1–4, 31. Then, we have chapters on understanding our musical brain and body for (and with) AI: Chaps. 5–9. These naturally connect with chapters discussing cognition and modelling thereof: Chapters 10–12 and 15. More technical chapters introduce a variety of applications ranging from improvisation and composition (Chaps. 14, 16–18), to performance (Chaps. 19, 27), orchestration (Chap. 20), notation (Chaps. 24, 25), studio production (Chaps. 13, 30) and even lyrics for popular music (Chap. 26). Other important applications of AI represented in this book are sound synthesis and signal processing (Chaps. 21–23), and musical robotics (Chaps. 28, 29).

The field is evolving faster than ever. AI is becoming so ubiquitous in our daily lives that the topic as we know it today is becoming diluted. It is being absorbed by other domains; almost every application of computing involves some form of AI in a way or another. One question that the AI and music community should certainly consider, however, is this: What is next? Has AI research reached the end of the road? Twenty years from now, what would the sequel book be about, if any?

The last three chapters (Chaps. 32–34) consider harnessing biology to develop living processors for 'not-so-Artificial' Intelligence systems and the potential of emerging quantum computing technology for music. The computers that our children will be using in 2050 are likely to be significantly different from the ones we are using today. What will be the impact of these on AI, music and indeed society?

I am grateful to Luc Steels for agreeing to write the foreword. Luc was director of Sony Computer Science Laboratory Paris at the time when I worked there as a member of his Origins of Language team. Luc is well respected internationally for his ground-breaking research on computational modelling of the origins and evolution of language, and robotics. Luc also enjoys composing opera[14] and has made pioneering contributions to the field of computer music in the 1980s. In the foreword, Luc touches upon the problem of making sense of musical meaning, a topic

[14]See *Casparo*, a tragi-comic opera in three acts by Luc Steels (music) and Oscar Villaroya Oliver (libretto): https://bit.ly/3obAHuH.

that is seldom discussed in AI and music research. It is an invitation to this research community to take the challenge.

I do not have words to express my gratitude to all contributors to this book. My heartfelt thanks to you all. You taught me a great deal about AI, but even more so about human intelligence, and above all, generosity. Thank you!

Plymouth, UK
October 2020

Eduardo Reck Miranda

Contents