HAL
open science

Jazz and Artificial Intelligence : From Presence to Traces

Marc Chemillier

To cite this version:

Marc Chemillier. Jazz and Artificial Intelligence : From Presence to Traces. In Bernard Lubat, Gérard Assayag, Marc Chemillier. Artisticiel / Cyber-Improvisations. Phonofaune, 2021, Dialogiques d'Uzeste., 2021. hal-03543103

HAL Id: hal-03543103

https://hal.science/hal-03543103v1

Submitted on 25 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Jazz and Artificial Intelligence
From Presence to Traces

Marc Chemillier
February 2020

In *Bernard Lubat, Gérard Assayag, Marc Chemillier. Artisticiel / Cyber-Improvisations. Phonofaune, 2021, Dialogiques d'Uzeste.*

All the improvisation softwares involved in this CD rely on the use of traces left by a musician (in this case Bernard Lubat). These traces can be collected by a microphone picking up the piano's sound. Musical notes can also be captured by an optical device such as the Moog Piano Bar, which is placed up against the piano's fall board to detect the movement of the keys (or MIDI systems, relying on the same principle, integrated in Yamaha Disklavier pianos). These traces can also come from a pitch tracker included in the software that extracts melodic patterns from the audio signal. This is the system we use to detect melodies in Bernard's voice. The resulting data obtained from those traces is then processed by three different softwares. OMax cuts the musical phrases played by the musician into fragments which are then recombined according to statistical properties. Djazz performs the same operation, but it bases itself on a temporal marking with even units, giving out a feeling of regular pulses. SoMax instantly creates a polyphonic co-improvisation that follows the live musician by drawing from a pre-learned corpus of models. As these softwares don't manipulate a raw audio signal, but the notes extracted from it, they can re-orchestrate phrases with a whole range of instruments accessible through a sampler. In the end, the computers don't produce a single sound that isn't derived, through one of the processes described above, from what Bernard played. This is intentional and results from choices we made in their design.

Thus, the exchanges exhibited in this disc between Bernard Lubat and improvisation softwares can be summed up as "Lubat and nothing but Lubat". But are the traces manipulated by the software really "100% Lubat", in the way some juice or yoghurt labels claim there are "real chunks of fruits inside"? When it snows overnight and footprints are found around the house, in the morning, we conjecture that an animal has passed by. It is no longer present because we can't see it, but it may very well still be hidden in a bush in the neighbourhood. These tracks left in the snow indicate the presence of the animal although it is no longer strictly there.

There is something paradoxical about wanting to improvise jazz using the traces left by a musician in the sense that they are, in a way, detached from his body. Because jazz is first and foremost a matter of body and bodily presence. The physical involvement of jazz musicians is considerable and is expressed in many ways that belong to the folklore of the genre: the grunts of Elvin Jones, Erroll Garner or Keith Jarrett (we can hear Bernard grunting in the Philadelphia recording), Cecil Taylor's drops of sweat, Art Blakey hanging out his tongue, etc. Moreover, the relationship between the musicians and the audience also takes the form of various bodily manifestations (clapping, "come on" and "one more time" interjections, dancing). This participatory relationship with the public is intrinsically linked to the ritual context in which music was practiced in African societies, where most of the ancestors of African-American musicians came from.

In jazz tradition, the music was played at a given moment, in a given place, with real musicians, for a flesh-and-bone audience. During our ImproTech Workshop in Philadelphia in December 2017, media theorist and machine improvisation specialist Bob Ostertag gave a brilliant lecture on the subject of improvisation. He emphasized, in particular, the essential nature of presence in defining what improvisation is: "What improvisation means is that a human still needs to be there. The meaning of improvisation becomes that it's the last claim humans have in the production of music." Which is the reason why jazz record collectors have long held the practice of closely reading any

liner notes and credits printed on the record's sleeve or booklet (who plays which instrument, when and where?). On the subject of what is regarded as the "classical" jazz era, great musical reviewers such as Hugues Panassié, could identify by ear the musicians of a recording session by analysing their technique, their touch, their phrasing and their sound, and could, if necessary, rectify a mistake on the cover. It should be noted that this practice tends to go out of fashion because most streaming platforms today hardly display liner notes anymore (Qobuz is a bit more comprehensive than Spotify or Deezer). Indeed, jazz lovers interested in a recording made in a specific year in this music's great tradition might find it difficult to find the exact version of the track they're looking for.

A recorded trace is always a reduction. This is a never-ending debate amongst jazz aficionados faced with the paradox according to which the studio recording is most often considered as an inferior version of the live performance, despite the fact that it's precisely those studio recordings that introduced most of these music lovers to jazz. But the reductive nature of the trace produced by audio technology goes far beyond the distinction between studio recordings and live performances. Technology is radically and profoundly transforming the practice of the concert itself. Bernard Lubat has often spoken to us about what he calls "technologies of representation" (technological means that come in between the artist on stage and the audience) to remind us of his attachment to a "dialectic between the living and the living" that should be established, according to him, during the concert between the person who plays the music and the one who listens to it. On this issue, he expands on the notion of "relation" developed by the Martinique poet Édouard Glissant (*Poetics of Relation,* translated by Betsy Wing, The University of Michigan Press, 1997). Still, the increased reliance on technology in the performing arts undoubtedly tends to reduce the relationship between musicians and spectators. The enormous scale of concerts (stadiums, arenas, etc.) made possible by sound technology is one of the most obvious symptoms of this tendency, and sometimes the musicians are only visible to the public thanks to the giant screens on which their image is projected.

In addition to sound amplification technology that makes it possible to reach ever larger audiences, other technologies contribute to reduce the number of musicians on stage, which also weakens relations. For example, the synthesizer—which originally had an expressiveness of its own—has gradually replaced the orchestra, allowing show organizers to cut costs. The DJ then replaced the band itself, and we see that today, in many music bars, the manager prefers to hire a DJ rather than a band of several musicians. Outside of jazz, in other musical genres, where the relationship between the artist and his audience is heavily influenced by the extra-musical phenomena of "stardomization" (variety, pop, rock), we are witnessing the development of practices such as playback that eliminate any form of relationship of a strictly musical nature. With the advent of disco at the end of the 1970s, the band Boney M built their career entirely on the use of playback. Of the four people on stage (one man and three women), the only one who really sang was Liz Mitchell and the others were just pretending. Today, Justin Bieber sings in playback in front of his fans, which, one can imagine, limits the risks of live performances in front of audiences of up to 20,000 people, as in Paris-Bercy concert hall.

Philippe Le Guern, Professor of Art Theory and Anthropology of the Contemporary World at the Rennes 2 University, France, is interested in the use of robots in art. He had invited Bernard Lubat, Gérard Assayag and myself to a meeting entitled "Artefacts sonores" (Sound artifacts) that he was organizing in June 2018 at the BPI in the Centre Pompidou, in Paris. He has published with Maël Guesdon a fascinating study in which he analyses the emblematic issue of holograms (*Où va la musique ?*, Presses des Mines, 2016). This technology is spreading today to bring deceased celebrities such as Whitney Houston back on stage. While the star passed away in 2012, she is expected to perform posthumously at Salle Pleyel, in Paris, in March 2020. But the most disturbing case lies with virtual star Hatsune Miku, who has no real identity of her own. This manga-esque

character appeared a little more than ten years ago endowed with a voice generated by the Vocaloid singing synthesis software, she now has about 2.5 million fans on her Facebook page and in 2013 she was the star of an opera entitled *The End* at the Châtelet theatre in Paris. Internet users can make Hatsune Miku sing using the Vocaloid software and make video clips using a 3D model of the star's body. The clips are then posted online, to be discussed and evaluated by the fan community. This participatory dimension is hailed on the dedicated forums as one of the newest and most fascinating aspects of the Hatsune Miku experience. But such a situation muddles up the question of presence. The archetypal link observed in stardom phenomena between the ubiquitous images of the stars shared by millions of fans, and the inaccessibility of the celebrity themselves, whose intimacy is protected by a security detail and protocols, is rendered completely obsolete in this case since there is no intimacy to protect.

The Crypton Future Media company that manages the hologram sells the simulated images and voice, but does so by imposing a top-down hierarchical structure to the organization of this community. The concerts must remain the model for all other images produced by the fans. They meet up at the official events organized by the firm, which simulates a shared emotional experience between the star and her audience, for example by making her cry on stage during a song. Videos of the concerts are then made available on the Internet and allow participants to keep track of the physical gathering of the community. Three researchers in cultural and arts studies from the University of Copenhagen have further analysed the economic model of Crypton Future Media in their article published in *Digital Creativity* (Volume 28, Issue 4, 2017). One of the original aspects of this model is to encourage fans to create music or video using the star's voice and image, which is very different from the approach observed with other famous characters such as Batman, whose merchandising is controlled by DC Comics, which holds the ownership rights, and mostly ignores the content generated by the fans. For these researchers Crypton Future Media's original IP strategy has to be considered within the larger context of the attention economy, which refers to a market system in which supply is abundant and therefore economically devalued (for example music on the Internet), while the resources that are scarce and sought-after become the time and attention of consumers.

The rise of false identities or "false presences" in the form of digital avatars is not an isolated phenomenon. On the contrary, they are extensively used online, in many domains. For example, interconnected websites known as "link farms" will point towards a targeted website for the purpose of improving its rank in search engines. We are also witnessing the rise of fake reviews of hotels, products and businesses written by cheap micro-workers in India or Madagascar who are hired to add rating stars on online products, making people believe that real customers have expressed real opinions. David Chavalarias, a researcher at the Centre for Social Analysis and Mathematics at EHESS (CAMS), where the improvisation software Djazz is developed, is a specialist on the subject of online information manipulation in the political field and one of the designers of an analysis tool called Politoscope. His team has obtained spectacular results on the use of the orchestrated disinformation technique known as astroturfing (see his paper on the web site AOC "Analyse Opinion Critique", 7 November 2018). It is a set of practices creating the illusion of a massive and spontaneous grass-root support in the political or cultural field. These techniques can be performed manually, when a group coordinates its actions to artificially boost a specific view online. Many activist groups join in such coordinated campaigns. But they can also rely on algorithms, using semi-autonomous "social bots" or when bigger players get involved (such as nation-states) through the use of a sophisticated "sock puppets" software allowing a single person to manage several hundred virtual identities, each with an active social media account (like Facebook or Instagram, etc.) with profile pictures and publications in order to make people believe they are interacting with real accounts. By mapping and analysing activist activity on Twitter, the Politoscope has demonstrated that groups from the American far right have used some of these

techniques to interfere with the 2017 French presidential elections in an attempt to promote Marine Le Pen's candidacy.

The problem of information manipulation was considered enough of a concern by the Ministry of Europe and Foreign Affairs and the Ministry of the Armed Forces that they commissioned a report in August 2018. The report states that the risk is not to change voters' opinions, but rather to spread doubt and confusion among them to encourage them to take action, for example to go out and vote ("Information Manipulation, a Challenge for our Democracies", in French, p. 25). We know that, in democratic elections, the result often lies within the mass of undecided voters, and that it is enough to convince a small percentage of voters with these messages to win. We should not exaggerate the harmful potential of digital technologies in this context. Propaganda has always existed, even in Tintin's adventures in *Land of Black Gold* (Methuen Publishing) where the Emir Ben Kalish Ezab used a plane to drop leaflets over the Bab El Ehr camp. But one of the strengths of digital technology in such political campaigns is that it allows "micro-targeting". This is exactly what the Californian company Cambridge Analytica was doing when it was caught in the midst of a scandal that broke out in 2018. It was revealed that it was using the personal data of 87 million Internet users to create highly accurate individual psychological profiles in order to target voters. In the face of such abuses of digital technology, however, we should not throw the baby out with the bathwater. Before giving birth to these practices, which are becoming industrialized and represent a real threat, the web was first adopted by simple activists. Fabien Granjon, a sociologist who for several years has been conducting extensive work on what he calls the "cultural front of popular resistance" deployed by Bernard Lubat in Uzeste, was one of the first to analyse the grass-root adoption of the web by social movements in his book *L'Internet militant* published in 2001 (Apogée Publisher). The widespread circulation of ideas on the Web is on the one hand stimulated by the activities of grassroot cyber-activists studied by Fabien Granjon and on the other hand blurred by the information manipulation campaigns described by David Chavalarias. We must tread carefully, because by attempting to clean it up too diligently we might risk shutting down the spaces of freedom introduced by the cyberspace in the first place.

One of the characteristics of the digital footprints left on the Web is that they can be quantified: for example, we can count the clicks of an Internet user. The notions of measurement and metrics have undergone considerable development since the neo-liberal policies of the 1980s have systematized their use in the field of public policies, alongside with the generalization of social and economic indicators. For researchers such as Gérard Assayag and myself, the use of quantified measures to evaluate our work takes the form of what is known as "bibliometrics". It measures the number of times a scientific article has been cited by others. As most publications are now available online, it is becoming possible to use algorithms to search the web to count all the mentions made by researchers when they cite articles by their colleagues. The idea is that the more a researcher is cited by others, the more these works can be considered of interest. Rating agencies have been developed to perform these calculations and their results weigh more and more on the career management of researchers (recruitment, promotion, etc.). In the book *Derrière les grilles* published in 2014 by Mille et une nuits, Barbara Cassin and Roland Gori vehemently protest against this system's absurd excesses. One of the most visible aberrations is the fact that it pays no attention to the content of the articles, i.e. what researchers actually write. The authors also denounce the perverse tendency of the system to function as a "self-fulfilling prophecy": the novelty is, by essence, not already known in advance and is therefore not cited, so that the mechanisms involved can only measure what is known, and the decision-makers who rely on them can only encourage and financially support what is already carved out and catalogued. As the authors point out, in such procedures invention is the great absentee.

The exploitation of the digital footprints left by Internet users has taken a form that is now enjoying a resounding success with the development of recommendation algorithms. The footprints left by

users on the web (which refer to a real activity which is tracked by algorithms, not the fake traces resulting from the astroturfing we mentioned above) are used to try to predict the future from the past. This principle is applied in a wide variety of situations, for example with the word suggestion on our mobile phones, which works by calculating the degree of likeness of the typed word with words stored in its memory, especially words that have been typed previously. Dominique Cardon, an Internet sociologist, has studied the recommendation mechanisms implemented in what is known as "collaborative filtering": users are shown things that have been bought by people who have previously bought the same thing as them. Online music distribution platforms such as Deezer or Spotify have made a cornerstone of their recommendation tools and they sometimes allow the music lover to make real discoveries. A parallel could be made with the serendipity of public libraries where you discover a book because it was on the same shelf as another book you borrowed. However, recommendation techniques carry the risk of trapping users within the limits of their own habits since they lead them to "adjust their desires to the regularity of their practices", as Dominique Cardon points out in his book *À quoi rêvent les algorithmes?* (Le Seuil, 2015, p. 88). More generally, they enforce a certain order of things which we wouldn't like to see becoming an imposed norm, without any possible recourse. It is a quandary inherent to all automatic modes that users should be able to disengage at any time in order to "switch into manual mode". All mobile phone users have once experienced refractory keyboard suggestion mechanisms that refuse to let them write the word they are desperately trying to type. Another source of concern is that those mechanisms are based on the footprints we leave as we navigate through digital interfaces. Digital advertising companies thus have an incentive to disseminate cookies in our devices and computers in order to collect these navigational traces. Such practices often operate at the limit of legality and represent a compelling threat to our digital privacy and the respect of our personal data.

The improvisation software we have developed is based on a similar principle, which consists of "predicting" the future from the past. The computer-calculated musical phrases are, in a way, an extension of Bernard Lubat, and his playing is treated as a set of traces collected by the algorithms. This process belongs to what is known as machine learning, which is a field of study of artificial intelligence. This discipline, which first appeared in the mid-1950s, has evolved over time, adapting to the problems at hand and the techniques available to solve them, but its ambition since the beginning has been to accomplish tasks normally reserved for humans. Musical creation belongs to such domains and, as such, we can speak of artificial creativity. Since the 1980s, the development of this field has resulted in the rise of systems based on trace collection and machine learning to the detriment of systems based on explicit rules. George Lewis, a musician from the AACM (Association for the Advancement of Creative Musicians), was one of the pioneers of computer improvisation by making computer programs interact with himself playing the trombone or with other great jazz figures such as saxophonist Roscoe Mitchell, another member of the AACM. He designed his first systems in the late 1970s by implementing sets of rules that formalized a musician's behaviour when improvising.

In the case of the artificial creativity software we are working on, our goal is not to predict entirely the future from the traces of the past (this would eliminate the idea of creativity). Rather, it's about producing things that look like the ones that have been recorded, while also inviting the system to introduce surprise and invention. This is made possible by the presence of random choices in the recombination process. In the end, what interests us most is not what the machine produces, but rather its confrontation with the musician himself. To be more precise, we should add that the confrontation involves in fact the machine, the musician and the two computer operators who set the parameters for the computations. Because the results generated by improvisation software also depend on human intervention, and therefore on the presence of real people to operate them. In the end, this kind of improvisation software is above all the occasion for an encounter between three people, Bernard Lubat, Gérard Assayag and myself, certainly mediated by computer devices, but in which the idea of presence is as important as if we were practicing jazz "the old-fashioned way",

that is to say without machines. From this point of view, the work we have been conducting for nearly twenty years with Bernard Lubat has been a constant source of wonder at the richness of his abundant feedback about this confrontation. It was apparent in his comments at the end of the listening session on 24 February 2020 during which we selected the recordings for this disc: “It gives me an impulse, it changes the way I play the piano. I realize that I’m navigating towards the worst. It cures my fear, we’re always afraid of displeasing, we’re always afraid of ourselves.”

Addendum – It is April 2020. Nearly half of humanity is currently confined in an attempt to curb the coronavirus epidemic. Performing artists are hit hard by this health crisis and the resulting wave of cancellations and closed venues. In such circumstances, the notion of presence takes on a singular meaning. Musicians are exploring different alternatives using new technologies as a means to recreate a sense of connection. Pianist Chick Corea livestreams himself every day on his Facebook page, from his home, improvising or playing scores on the piano (Chick’s Workshop). At times he gets up, walks up to the camera to see if people are connected and, if so, sends them a friendly message. We are witnessing a rise in the use of split screen video editing, displaying a mosaic of individual screens, enabling whole orchestras not to play live but to mix together parts that are recorded remotely and separately (the French National Orchestra’s rendition of Ravel’s *Bolero* drew over 2.4 million views on YouTube in March 2020). Listeners are also attracted in increasing numbers to social listening services on music streaming platforms which enable people to listen to the same song simultaneously over the Internet. Technology is enabling these virtual relationships mediated by the Internet through video and social networks. It is difficult to say what will happen after the coronavirus crisis. We will no doubt return to a traditional musical relationship, but many music venues will not reopen and the world of performing artists will suffer heavy losses in this crisis. Will musical practices based on a “virtual” mode of presence that is entirely dependent on technology also continue to exist, and if so, what will the long-term consequences be from an aesthetic, psychological, social and environmental point of view?