The Routledge Companion to Jazz Studies

Nicholas Gebhardt, Nichole Rustin-Paschal, Tony Whyton

Jazz on Radio

Publication details

Tim Wall
Published online on: 06 Dec 2018

How to cite :- Tim Wall. 06 Dec 2018, Jazz on Radio from: The Routledge Companion to Jazz Studies
Routledge
Accessed on: 20 Aug 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
As for jazz and computers, the question is what they could possibly have to do with each other. The values and processes of the one—an art of spontaneous, inspired collective improvisation, evolving through an oral tradition, born of African-American culture but accepting creative individuals by way of true meritocracy—seem antithetical to the other. . . . I can imagine computers delivering goods and services on a global scale, much more efficaciously than ever before. But I have a hard time conceding they’ll ever provide us with the sustaining reflections of jazz.

Howard Mandel (1997)

What makes the two things work together is that even though they seem so incompatible the premise of free jazz is that you improvise with whatever is in your environment. If it’s a machine, you interact with that.

Matthew Shipp (in Nicholson 2003)

In his 1935 essay “The Work of Art in the Age of Mechanical Reproduction,” Walter Benjamin interrogated how mechanical reproduction removes the “aura” surrounding art by confounding its unique existence in time and space, thereby altering perceptions of its originality, authenticity, and embeddedness in tradition. In addition, he was concerned with how new technologies shift or augment human perception, cultural production, and social and political participation. For Benjamin, like many of his Frankfurt School colleagues, new modes of appreciation and engagement arise in tension with new modes of deception and distraction. This chapter extends these questions into the contemporary era by interrogating the art of musical improvisation, often considered to be an essential aspect of modern jazz that provides it with an “aura,” in settings that also involve computational participation.

Jazz music evolved in tandem with the development of mechanical and electronic sound recording, and with the evolution of sound reinforcement, transmission, and production technologies (Myers 2012). More overt electronic sound manipulation has also been part of the jazz soundscape since at least the late 1950s, but until quite recently computational approaches were too slow and too cumbersome to engage effectively with musicians in the live situation, unable to offer the sense of spontaneity, discovery, and interaction cherished by jazz performers and listeners. The term computer music may still evoke room-sized contraptions from the 1950s slowly churning out grist for an elite composer’s mill, or, more recently, someone hunched over a laptop making barely perceptible movements on stage (perhaps checking email? as the well-known joke goes), neither of which seems easy to reconcile with the emblematic intensity and spontaneity of a jazz
improviser. Increasingly, however, computers—in a variety of guises and with a diverse array of “abilities”—share the stage with human performers, affording not only an expanded sonic palette but also the possibility to transform auditory space and time, to distribute creativity widely across persons, things, and locations, and to invite new musical actors, new forms of musical agency, and new modes of musical perception and production into the artistic experience.

This essay explores this post-millennial moment by tracing a variety of approaches to, and attitudes about, human–computer interaction, specifically in jazz and improvised music. It does not pretend to chronicle all of the diverse activities of contemporary improvisers who invite computational participation into their musical practice, by now an unwieldy task. Instead, it focuses on a few examples of computerized systems that present themselves as semi-autonomous agents, having the capacity to invent, provoke, and respond, rather than as “instrumental” extensions of a human performer or as “proxies” for a human composer, with the hope that this may provide some insight into larger trends and salient issues.

For instance, what are the repercussions of exploring musical improvisation through an algorithmic, computational, and combinatorial lens? Do approaches such as these invite us to rethink deeply held assumptions about creativity? Do they encourage novelty and innovation, or might they actually run the risk of undercutting difference by emphasizing and redistributing what is already statistically common? Can we conceive of improvisations as mediated by the skills, bodies, and desires of all the participating actors and agents, including non-human ones? How do artists and researchers reconcile the in-time, phenomenologically rich, and open-ended dimensions of actually doing musical improvisation with the over-time, reflective, and symbolically mediated dimensions of designing non-human improvising agents? In short, in what ways does computational modeling simultaneously elucidate, challenge, and perpetuate normative conceptions of improvisation?

In this chapter, I argue that the details surrounding the implementation of a given improvising system can shed light on how designers conceive of improvisation and interactive processes, while also illuminating how performers perform their personal identities and ideologies through specific strategies of interacting with these systems. Like Benjamin, I am also interested more broadly in changing modes of human perception and participation, and in that light, I hope that this chapter, focused as it is on the aesthetic and conceptual work that musical improvisation can do in the computer age, might contribute not only to the “New Jazz Studies” but also to the wider discourse on new media and technocultural studies.

**Strike Up the Virtual Band**

The notion of a “virtual” band in jazz has arguably been with us since the introduction of “Music Minus One” and Jamey Aebersold “Play-A-Long” recordings. None of these, however, could be called interactive, in the sense that they are capable of responding to human performers in the course of performance or of using input information derived from a musical performance to shape compositional decisions. Every time one puts on a record to play along, one hears the recorded musicians develop the performance in exactly the same way. These can provide useful practice tools for developing musicians who may not have access to—or may not wish to burden—a live rhythm section (although this likely has also contributed to the reification of the soloist in jazz pedagogy). The opportunity to play alongside seasoned jazz musicians (“Is that Ron Carter on bass?”) undoubtedly contributes to the success of this format as well. The technological medium of conventional audio recording, however, does not provide interactivity or an ongoing sense of musical surprise.

More recent computerized accompaniment programs, such as *Band-in-a-Box* or *iReal Pro*, offer additional possibilities, usually by allowing users to decide upon aspects of the performance
The Art of Improvisation

beforehand (for example, tempo, key, meter, or style), after which the system generates an appropriate accompaniment. Once generated, however, this accompaniment functions more or less like a play along record (although Ron Carter will likely have been replaced by Midi notes triggering sampled bass sounds). Again, these software tools can be useful practice aids (for example, the current version of iReal Pro offers the option for automatic tempo increase and transposition at every repeat of the song form), but without any listening or learning functionality they cannot offer real time musical interaction.

François Pachet, the director of Sony Computer Science Laboratory Paris, is a jazz guitarist who has developed several systems for interactive musical performance. His projects titled VirtualBand, Virtuoso, and Continuator are the most relevant to the current discussion. VirtualBand is the outcome of Pachet’s desire to make a truly interactive play along system (Moreira, Roy, and Pachet 2013). Like other machine learning applications, VirtualBand has a “training” period during which musicians on different instruments (for example, guitar, piano, bass, drums) are recorded in studio situations playing in a variety of styles. These recordings are later analyzed by the system using signal processing techniques (primarily based on music information retrieval research) that segments and tags different parts of their performances and organizes them into a style database. To use VirtualBand, one selects a song structure (essentially information from a jazz lead sheet such as harmony, meter, and form) and the number and type of accompanying musicians. The important difference, according to Pachet, between a standard play along and VirtualBand is that his system is designed to respond to the dynamics and musical development of a human soloist as the performance unfolds. For instance, as a user builds a solo in intensity (perhaps by increasing volume and/or density of notes), the system detects this change and is designed to respond accordingly.

Continuator is a related project that offers a kind of pedagogical scaffolding and performative interactivity to jazz soloists. As its name suggests, Continuator is designed to continue the musical ideas being developed by a human partner in real time in a stylistically appropriate fashion. It is designed to “learn” an individual musician’s performing style in real time so that it can engage that musician in a type of call-and-response playing. An earlier project with a similar goal designed by Belinda Thom was humorously titled Band-out-of-the-Box. Videos of Continuator in operation demonstrate engaging moments of musical “trading” as the system quickly builds operational representations of the current style being explored by its human counterpart, and one video in particular pits the system against a skilled improvising pianist in a musical “Turing test” that, at least according to the comments provided for the video by Pachet’s lab, had two music critics deciding largely in favor of the computer.1

A third system being developed in Pachet’s lab, Virtuoso, generates a highly virtuosic style of improvisation by taking input from a human operator using a Nintendo Wii remote with a Nunchuck controller to steer the automated production of individual notes and longer, rapid sequences or phrases, all of which can be constrained by the user choosing a default scale (such as diatonic, chromatic, blues, or diminished). The system’s virtuosity easily exceeds the technical abilities of even the most rapid-fire instrumentalists, sounding in practice something like a hard rock or bebop musician on steroids.

A crucial point to make is that each of these three systems requires that a considerable amount of standard jazz theory and conventions be built in—either to the software or to the process of training it—for them to function in expected ways. For instance, the computer drummer in VirtualBand switches from playing gentle brushes on the snare drum, to using sticks on the high-hat cymbal with occasional “bombs” on the snare, to eventually keeping time on the ride cymbal and incorporating more pronounced hits on the toms, all in response to the perceived energy of a soloist. Similarly, the bass player in the “virtual band” switches from walking in 2 to walking in 4 as the perceived intensity and density of input from the soloist increases. These are conventions of mainstream jazz accompaniment but are certainly not the only way that drums and bass can
increase their dynamism (and arguably some of the most interactive mainstream jazz does not always conform to these conventions, such as the Bill Evans trio).

Likewise, while the Continuator system appears to offer a promising way for improvising musicians to engage a computer system that quickly learns, adapts to, and extends one’s own ideas, it, too, has certain cultural conventions “baked” into its design. The most obvious moments of joy in the video documentation are when the musicians interacting with Continuator hear their own ideas coming back to them in slightly modified form. This kind of cat-and-mouse playing, however, is frowned upon in some improvised music circles, and one may tire quickly of a system that constantly mirrors one’s own behavior or requires continual prodding in order for new musical ideas to emerge.

Lastly, the Virtuoso system is certainly capable of generating fast and intricate melodic “lines,” under the “steering” control of a human operator, but to do so, it conforms explicitly to underlying chord-scale conventions that betray a particular historical, cultural, and institutional bias (Wilf 2014). As was alluded to at the outset, these implicit biases designed into computational systems may serve to further codify and disseminate standardized practices in ways that could ultimately discourage real novelty and innovation.

He Has Heard

Shimon is a robotic marimba player capable of improvising complex jazz-inspired melodies and chordal accompaniment. Currently, it uses four mallets, each attached to an independently moving mechanical arm, which are in turn connected to a large mechanical torso with a protruding head and a single large eye. This makes Shimon vaguely resemble Luxo, Jr., the animated lamp that serves as the “i” in the logo for Pixar films, which was an early design inspiration.

According to Gil Weinberg, one of Shimon’s developers at the Georgia Institute of Technology, the term “robotic musicianship” may strike some as an oxymoron, since to “play like a robot” is a phrase most commonly used to critique a human performer who is unable to add the requisite musicality and sensitivity to “lift” the notes off the printed page, offering only a “mechanical” reading of the score (Bretan and Weinberg 2016, 100). For Shimon’s team of researchers, however, the pursuit of robotic musicianship is an interdisciplinary scientific and artistic endeavor. It has an engineering side, which involves the study and construction of physical systems that generate sound through mechanical means (“musical mechatronics”), and a computational side, which focuses on developing algorithms and cognitive models representative of various aspects of music perception, composition, performance, and theory. Mason Bretan, a key member of Shimon’s research team, explained in a phone interview how Shimon’s playing is shaped by the material affordances and constraints of its mechanical design:

It has a representation of some musical language or some musical knowledge, but then on top of that, when it makes its decisions it thinks about: “I have four arms and they can move in this particular way, so how can I achieve this particular goal.” So something as simple as playing a motif that goes up in pitch, Shimon really has to think about its body in order to do that. So its body influences what note it might start on within the motif, and where it is going to end, and what it is going to do in the middle, and how fast it can go is all dependent on its body.

(Bretan 2015)

Shimon has also been programmed with a repertoire of listening techniques and “ancillary” movements—those not directly related to sound production—that are designed to assist with musical coordination and to provide a more engaging performance. For instance, in addition to
its anthropomorphizing value, Shimon’s “eye” serves as a visual tracking system that allows it to notice movements made by its co-performers and use this information to attempt to synchronize certain musical behaviors, similar to how human performers can use their own eyes to track Shimon’s body movements to assist with musical coordination or anticipation.

For audiences, seeing Shimon wildly swinging mallets, twisting its neck to follow its co-performers, and bobbing its head to the beat, appears to influence the perception of the robot and the music it creates. In one early experiment, listeners were either provided audio-only or audio-visual recordings of Shimon performing with a human musician. With the audiovisual recordings, participants consistently reported higher levels of appreciation of how well the robot played, how much it played like a human would, and how responsive it was to the human performer and the human performer to it (Hoffman and Weinberg 2015). Shimon is a Biblical name that means “he has heard.” It appears that being able to see the robot’s physical embodiment and engaging movements contributes greatly to a sense that it is listening well.

That being said, Bretan, a jazz percussionist and pianist himself, admitted to me that his early encounters with Shimon felt like performing with a really amateur musician who needed a strong and clear articulation of the beat in order to stay synchronized.

It has improved a bit, and now there are moments where it feels like I am interacting with a person . . . [but] it is hard to get Shimon to progress the music, the interactions . . . to create a beginning, and end form, a sort of higher-level structure. Shimon makes a really nice call-and-response demo, but it needs to have a better understanding of higher-level structures . . . it has to be able to tell a story.

(Bretan 2015)

The Georgia Tech research team has frequently articulated an overriding goal of the project is to create a system that can “listen like a human, but improvise like a machine” (Murphy 2016). Having the two additional independent arms already gives Shimon the possibility of playing chordal and melodic structures on the marimba that are spaced beyond what even a skilled four-mallet human performer could achieve. This raises the question of whether Shimon’s specific mechanical abilities and constraints might lead to its own idiosyncratic style of improvising. Eitan Wilf, who conducted extensive ethnographic research at Shimon’s lab, quotes one of the researchers describing Shimon’s improvising style:

You see, [Shimon] has its own style because of the arm movements and the limitations. You’d hear the beginning of a natural run [i.e., a phrase that consists of notes adjacent to one another] and then suddenly a note would go up in the octave—you’d hear some note being played by a different arm in a different octave because the first arm is not fast enough to play it so the other arm would compensate for it. And I think that’s unique to [Shimon].

(Wilf 2013, 728)

One might argue, however, that there is a certain disconnect between programming Shimon with human music—melodic, harmonic, and rhythmic structures that are indelibly stamped by the physicality of human performers and human instruments—in order to have it produce machine music.

In the more theoretical portion of his work, Wilf writes compellingly about both the material constraints and the normative ideals that inform computational and mechanical approaches to modeling improvisation. For example, Shimon’s research team has recently explored ways of imbuing the software with an ability to mix improvising styles by combining aspects of the styles
of historically important improvisers, as derived through statistical analysis of transcriptions of their improvisations. Wilf writes about a session in which Shimon was tasked with improvising with a mix of 33.3 percent the style of Miles Davis, 33.3 percent the style of Charlie Parker, and 33.3 percent the style of the player improvising with Shimon, which Shimon could learn in real time. According to Wilf, not only does a residual style still pervade Shimon’s playing regardless of which mix of styles it is asked to perform, at least in part a result of its physical embodiment, but this tactic of quantifying creativity, dividing it into building blocks, and recombining them, betrays many normative beliefs and practices (ibid., 719). What is deemed important in music is invariably culturally inflected, and Wilf argues that Western music analysis has a long history of emphasizing an underlying parametric paradigm of creativity, which can now, through computational mediation, be taken to its logical extreme.

I also worry that by dividing mechanical from computational dimensions, and listening from generative procedures, this type of research subscribes to binaries that are unsustainable. When I asked Bretan if Shimon was listening to itself, he laughed and admitted that they “try to isolate Shimon from being able to listen to itself because in order to get a clean audio signal they need it to focus only on the other sounds around it” (Bretan 2015). This, to me, highlights a significant shortcoming, in that improvisers listen intently to their own sounds, along with those of others, and often explore new musical directions based on intimate forms of auditory and haptic feedback and feedforward provided by the tight coupling of player and instrument, in addition to that between the various musicians.

At its best, however, this research does offer substantial empirical insight into issues of timing, anticipation, expression, mechanical dexterity, and social interaction—at least in terms of expressive behaviors and musical cueing—that are central to music making and to human behavior more generally. At times, however, the researchers’ aspirations seem to get the best of them. For instance, Weinberg and Breton write that Shimon will inspire humans to “invent new genres, expand virtuosity, and bring musical expression and creativity to uncharted domains” (Bretan and Weinberg 2016,102), a type of language not only filled with teleology and hyperbole, but also seemingly unaware of the ways in which it is culturally situated and arguably gendered as well, in that it appears to reflect stereotypically “masculine” experiences and prejudices.

**Intimacy and Opacity**

Maxine is a software-based improvising system designed for more experimental forms of free improvisation created by Ritwik Banerji, an ethnomusicologist from the University of California Berkeley, who is currently a fellow at Free University of Berlin. In intriguing ways, Banerji envisions Maxine as both a performance system and as a tool for ethnographic investigation of the improvisation community (Banerji 2016). I begin with a brief discussion of the technical aspects of the system before exploring the ethnomusicological dimensions of Banerji’s project.

Maxine is realized using Ableton Live software and Max/MSP, a programming language commonly used by computer musicians whose name Banerji borrowed and intentionally feminized. Maxine’s design uses multiple agents analyzing auditory input and controlling sonic output simultaneously. Many of these agents are linked to a pitch detector, the Max/MSP object called [pitch~], that, when presented with the complex and often pitchless sounds of free improvisation, usually fails. This means that Maxine’s output will correspond in general to the overall pacing and event-density heard in the playing of a human partner, but how and why particular correspondences are made can be difficult to deduce. For the sound producing agents, Banerji employs digital instruments and processes from Ableton Live that create “unusual” timbres, such as “metal percussion, synthesized versions of prepared or ‘extended’ guitar and piano techniques, a variety of synthesizers, and signal processing tools” (ibid.).
After experiments with early versions of the system, Banerji (a saxophonist himself) decided that Maxine’s output conformed too closely to the temporal and timbral developments of its human partners. He felt that Maxine required too much prodding to produce things of interest (a concern I raised with respect to Pachet’s Continuator above). To address this shortcoming, Banerji decided to add a type of cybernetic feedback to the system by positioning an additional microphone close to the system’s own speaker output (a marked contrast to Shimon’s design team, who intentionally avoid having it listen to its own playing). While it is conceivable that this setup would create runaway feedback, Banerji found that the unusual timbres he chose for Maxine’s output continued to confound its pitch detector, providing a greater sense of “mystery and individualism” to its output (ibid.). Somewhat fortuitously, Banerji realized that he could alter the distance from speaker to microphone as an additional variable affecting the system. For human partners who desire a more “aggressive” profile, Banerji positions Maxine’s microphone closer to its speaker, providing it with more independence. For those wishing for a more “sympathetic” response, he moves the microphone further away, causing its output to correspond more closely with theirs.

With these new possibilities in place, Banerji began soliciting post-performance feedback from Maxine’s musical partners for insights that might lead to further improvements in system design. Some players found Maxine to be “too meek, hesitant, or reserved in its interactive behavior,” unable to inspire their own improvising (ibid.). Others felt that Maxine was too “self-absorbed,” unable to “meet them halfway” (ibid.). Some felt Maxine needed to be more skilled at recognizing and (re)producing conventional musical features, such as meter, interval relationships, harmonic progression, standard phrasing, etc., while others wished for more unpredictability. Interestingly, Banerji found that opposing traits were sometimes desired by the same individual, depending on the context of the unfolding musical performance. After all, while musical “aggression” in one context might be desirable, in another it may be repugnant.

Many felt, in ways similar to Bretan’s concerns with Shimon, that computer systems have difficulty creating the impression of evolving musical form across longer expanses of time—of “telling a story.” One participant expressed that Maxine demonstrated “a frustrating inability to sustain the drama of the interaction” (ibid.). Others expressed that Maxine lacked the important ability to be able to switch between leader and follower roles, or to provide longer periods of either support or opposition in musically appropriate ways. Many wished that the system could produce musical materials to mesh better with their own, often expecting Maxine to emulate the style of human improvisers personally familiar to them.

Banerji soon realized that the feedback he was receiving offered valuable insight into the unspoken norms and cultural politics that emerge in scenes of musical improvisers, a kind of insight that is not easily attainable since free improvisers tend to avoid conveying these kinds of explicit preferences or critiques to one another, likely out of deference to their musical liberty. Since Maxine is a non-human agent, participants felt far more comfortable critiquing its performance. Through this process, Banerji also became more self-aware that his design decisions with Maxine betrayed his own cultural location, making the interactions between Maxine and its partners an opportunity to elucidate the relative distances between their cultural locations and his own.

Ultimately, Banerji realized that it is impossible to separate out an evaluation of Maxine’s performance provided by her human co-performers—and here I intentionally shift to using the feminine pronoun—from an evaluation of their own performance, in terms of how they interacted with the system. As an improvising “system,” Maxine’s performances do not rely on traditional notions of leader and follower, or reify ideas about instrumental virtuosity or creative teleology, which appear to be dominant motivators in the labs of Pachet and Weinberg, based on their system designs and published comments. Any assessments of musical quality in a performance with Maxine are embedded in a specific socio-cultural context and are mediated between the skills
and desires of all agents, including the human participants, Maxine and Banerji, and all of the programmers and designers behind Ableton Live and Max/MSP as well.

On a purely technical level, it can be argued that Maxine is a less complex computer system. It is not based on a massive database of transcribed solos or on complex “style modeling” methods. In fact, it does not attempt to represent musical knowledge in any way. Its primary feature extraction method (pitch tracking) is arguably poorly chosen, given the kind of musical “information” it will most likely encounter. Yet, because of these things, rather than despite them, it represents an overture toward a different paradigm for understanding and modeling collective improvisation.

Traditionally, artificial agents have been programmed to sense, represent, compute, and respond to an agent-independent pre-given world. Banerji’s approach, by contrast, is more in line with a cybernetic view that upholds a vision of the world as a place of continuing interlinked performances, what Andrew Pickering calls a performative ontology (Pickering 2010). A performative ontology insists that agents learn through an ongoing cycle of perturbation and compensation that resists the detour of knowledge through representation (see also Borgo 2016). Cybernetic systems are not programmed to seek goals in reference to predefined states of the world—goals which would necessarily be extrinsic rather than intrinsic—and their actions cannot be fully explained through linear cause-and-effect relationships. Perturbation, after all, contains the root word turbid, meaning opaque.

Banerji’s system does not have any explicit memory, learning, or predictive capability, which might be seen as a significant detriment, but when viewed as a cybernetic system, his interventions in the system (e.g., altering microphone placement or revising the number and behavior of software agents), along with how Maxine’s co-participants alter their musical behaviors to mesh better with hers, can all be viewed as forms of systemic learning. In this more cybernetic view, humans both configure—and are configured by—the technologies around us (Borgo and Kaiser 2010).

**Telling a Story**

In the same year that Benjamin wrote his well-known essay “Work of Art in the Age of Mechanical Reproduction,” he authored a less well-known essay titled “The Storyteller.” These essays deal with different technologies—cinema and photography in the former, and printing in the latter—and seemingly offer rather different perspectives on our engagements with technology: a difference that some have read as a change of heart in Benjamin’s own thinking. In “Work of Art,” Benjamin celebrates the potentially democratizing aspects of cinema and photography, as well as their ability to alter human perception and experience through new camera-angles and editing techniques such as slow-motion and close-up. In “The Storyteller,” he mourns the passing of storytelling as a lived, oral tradition under the impact of the printed word and the pre-digested nature of information conveyed through newsprint. Benjamin feared that storytelling was disappearing in modernity because face-to-face contact and living praxis were giving way to “information” as decontextualized and instrumentalized knowledge. Read together, however, these essays demonstrate less a change in heart than a call to be cognizant of the different ways that different technologies interact with culture, and of how the same technologies can lend themselves to diverse and divergent readings (see Hogg 2012).

Without a doubt, there have been significant strides made in artificial intelligence and machine learning in just the past few years, but consideration of social and cultural issues within these research communities remains rare. One obstacle is the simple fact that gaining the requisite knowledge in both the technical aspects of learning algorithms and in the critical theories of the contemporary social sciences and humanities is an imposing challenge. Media coverage of this research also tends to promote sensationalist views that favor hype over reality, using dystopian or utopian rhetoric in place of real critical engagement. The following, for instance, are recent...
headlines from articles covering the activities of the robotic musicianship research group that created Shimon: “A Four-Armed Robot Can Now Improvise Music as Well as Human Bandmates”; “This Robot Is a Better Jazz Player Than You”; and, perhaps with a sly reference to this trend, “Your New Robot Overlord Turns Out To Be A Pretty Good Marimba Player.”

Nick Collins, who has both technical skills in this domain and a desire to address relevant social and ethical issues, argues that these emerging musical agents are best understood as “projected intelligences” rather than artificial intelligences, since they result from a composer’s anticipation of the dynamics of a concert setting made manifest in programming code (Collins 2006). Collins admits in a slightly later publication, however, that “machine musicianship continues to advance, and machine learning techniques may undermine many certainties here” (Collins 2011).

One thing that does not change, however, is the ways in which specific technologies mark and often maintain specific cultural positions and presuppositions. For instance, Cathy O’Neil’s (2016) research suggests that sexism, racism, and other forms of discrimination are being built into machine learning algorithms simply because these already exist in the data that is being learned. Just as face recognition software trained on photos of people who are overwhelmingly white will have a harder time recognizing non-white faces, it stands to reason that computer improvisers trained on a heavy diet of Sonny Stitt and Michael Brecker will have more difficulty “understanding” Ornette Coleman and Albert Ayler. Learning algorithms are only as good as the data they are trained on (“garbage in, garbage out”), and data of this sort has a complex history.

Extracting only a “solo” voice from the complex, collaborative, and social environment that informs jazz musicking is already problematic, as is the reliance on parametric representations of jazz “vocabulary” as analyzed through transcription of recorded solos. While music may have some language-like qualities, it is not language, and its differences should not be minimized or ignored just so that researchers can conveniently borrow deep learning algorithms that have proven successful in that realm. Recognizing and predicting correlations between written text is not the same as developing spoken fluency (which involves the pragmatics and prosody of language use, among other things), nor would generating expert-level musical improvisations using the symbols of music notation be the same as performing and interacting fluently in a live context.

On the whole, researchers in artificial intelligence have realized that it is relatively easy to get computers to do formalized adult things (like play chess and do math—or perhaps even to play a generic form of bebop), but decidedly more difficult to get them to do things that come naturally to a small child (like play with a ball, identify a dog, develop a sense of humor, or entrain to a musical beat).

It is also worth noting that designing computers to learn how to emulate jazz—either from recorded or real-world examples—and then injecting that emulation back into the world ultimately affects the world that the system is intended to emulate. This is not simply a pedantic question of which style will be favored in the world of computer-based jazz. It is also a pressing question of diversity and inclusivity. Kate Crawford (2016) writes:

Like all technologies before it, artificial intelligence will reflect the values of its creators. So inclusivity matters—from who designs it to who sits on the company boards and which ethical perspectives are included. Otherwise, we risk constructing machine intelligence that mirrors a narrow and privileged vision of society, with its old, familiar biases and stereotypes.

To invoke Benjamin one final time, a story differs from information in that it allows—even demands—that listeners interpret, reflect, share, remember, and transform it in their own way—to make the story their own. Narrative, according to Benjamin, “achieves an amplitude that information lacks” (Benjamin 1968, 88). Information, seen only as raw data, denies wisdom,
shared experience, and the networks that forge community. In “The Work of Art in the Age of Mechanical Reproduction,” Benjamin interrogates not only the modern media technologies of cinema and photography but also psychoanalysis, which provided new understandings of the workings of the unconscious mind and promised a therapeutic method for bringing this material into our conscious awareness. The thread that connected these for Benjamin was their ability to provide the necessary critical distance for new understandings to emerge.

Working with computers has provided new ways to share and experience the social and psychological dimensions of our lives, but it may be the critical distance they can offer on our own creativity, our own humanity, and our own failings that will prove to be among their greatest contributions. In an interview (Nicholson 2003) exploring his approach to combining electronics and jazz improvisation, Matthew Ship remarked: “The machine is something that takes you outside yourself, but I’m actually finding the machine is allowing [me] to connect more inwardly with myself.” “It’s a kind of a paradox,” he mused, “but it’s really fascinating how it’s working out.” I am hopeful that insights like these, and the critical distance offered by the academic community as well, will continue to inform and propel the exploration of the art of improvisation in the age of computational participation.

Notes
2. Although Wilf anonymizes his sources in this article, in a personal communication with the author he confirmed that it is based on research done in the Shimon lab.

References