Performing with Technology: Lessons Learned from the GenJam Project

John A. Biles

School of Interactive Games and Media
Rochester Institute of Technology
http://igm.rit.edu/~jabics/
jabics@rit.edu

Abstract
The author has been performing with GenJam, the Genetic Jammer, for nearly 20 years and has accumulated a wealth of experiences in performing live jazz with technology. This paper presents a discussion of the use of technology in jazz, both from the performer’s and from the audience’s perspective, and it proposes a classification scheme for live performance that is geared to mainstream performing situations.

Introduction
GenJam, the Genetic Jammer, is a real-time, interactive performance system that uses evolutionary computation to evolve melodic material collaboratively with its human sideman as it improvises on straight-ahead jazz tunes (Biles 2013, Biles 2007, Biles 1994). This paper focuses on performance aspects of GenJam and compares GenJam performances to other jazz performances that use technology in order to motivate a classification scheme that provides insight into how technology is used by live performers and perceived by audiences.

The word technology is an ambiguous term. In this paper I use technology to refer to electronic/digital technology that impacts the aural content of a performance. This includes synthesizers, effects boxes, MIDI systems, and computers. I exclude acoustic technology (traditional instruments), sound reinforcement, and video. I ignore most electroacoustic instruments and controllers because my focus is mainstream jazz, which rarely uses such devices.

My playing with GenJam is unusual for a computer music system in that we perform in an established, traditional musical genre: straight-ahead jazz. I’m not trying to extend the boundaries of jazz or improvisation, and I’m not trying to challenge an audience’s conception of what jazz, or music in general, is or could be. I am trying to have fun creating and playing music that I enjoy and to connect with a mainstream audience. If there is any agenda for the technology that GenJam represents, it is to demonstrate that technology can be creative, collaborative and engaging.

This paper, then, focuses on performance issues that I have experienced in my nearly 20 years of GenJam gigs. To provide a context, I’ll describe some recent experiences I had as an audience member at a major jazz festival, and I’ll discuss how performance gestures, interactivity, and spontaneity come together, both for live jazz performers and for their audiences.

The audience perspective is important, yet often overlooked. In electroacoustic music, live performance is well established, and considerable research on novel performer interfaces is ongoing, as illustrated by the work presented at the annual New Interfaces for Musical Expression conference (NIME). This work, however, almost always takes the performer’s perspective and almost never the audience’s. Aside from occasional calls to consider the audience (Schloss 2003, Reeves 2005), or the rare in-depth discussion of performances that include audiences as stakeholders (Bower 2003), the audience seems to have to fend for itself in most electroacoustic performances.

By performing unabashedly straight-ahead jazz in my GenJam gigs, I seek to engage a mainstream audience with a hopefully familiar genre. When discussing my use of technology with GenJam, then, I compare GenJam performances to jazz performances that happen to use technology, instead of technology performances that might be called jazz. This brings us to the best nine days of the year!

Best Nine Days of the Year
The Rochester International Jazz Festival (RIJF) marked its 12th anniversary last June 21-29 with 312 concerts by over 1200 musicians at 19 different venues in downtown Rochester, NY (XRIJF 2013). One great feature of the festival is the Club Pass, which for a very reasonable price al-
lows the holder to attend as many as five sets every night of the festival. For a jazz lover, like myself, it’s simply the best nine days of the year. It’s also a fascinating opportunity to hear a stunning variety of jazz from all over the world geographically and all over the map aesthetically. At last summer’s festival, I took particular note of the uses of technology in the performances I attended, which provided some insight into how technology is used by jazz performers and received by audiences. I’ll detail two specific performances out of several that were instructive.

THIEFS (THEIFS 2013) is a trio consisting of tenor sax, upright bass and drums, a very traditional instrumentation. However, all three musicians use technology extensively to augment their performances. The drummer/vocalist uses a drum pad to trigger audio samples ranging from electronic percussion hits to phrase-length samples, and the bass player uses synthesis, distortion and reverb effects to color his lines. The tenor sax player, however, was the most interesting. Christophe Panzani had 11 electronic effect boxes arranged at his feet, bristling with pedals, stomp buttons and knobs. The effects included reverb, modulation/synthesis, a harmonizer and a loop delay, all configured in an effects loop familiar to many guitar players.

Panzani’s performance gestures on stage were noteworthy. On one piece he set the effects parameters on various boxes before the tune and just played his sax, which resulted in a focused and nuanced performance. During the other tunes, he tweaked the parameters on multiple boxes, which required that he tap his foot, operate foot pedals, and turn knobs, usually while playing his tenor sax. When he did serious foot tapping, he tended to play substantially fewer notes on the sax, both because he was listening to the output to fine tune the sound, and because he had to execute a lot of taps to get to the desired mode on some effect boxes.

When Panzani tweaked knobs, he literally knelt on one knee, rested the sax on the other knee, playing it with his left hand as he turned knobs with the right. Needless to say, this inhibited what he could produce on the tenor, both because he had a limited set of notes he could play left-handed, and because he was navigating a non-trivial and unmusical interface on the boxes.

The result for me was a bit off-putting. The sounds he was getting with the effects were fun and added a dimension to his improvisation, but the loss of focus on his melodic lines was frustrating, and the visual aspect of the performance was distracting. At one point, he was tweaking a knob on a harmonizer to alternate between a major third and a minor third applied to notes in a phrase loop he had built up earlier in his solo with the loop delay. The result was interesting to hear, but less than compelling to see since he was simply kneeling and turning a knob back and forth for several seconds.

I suppose my negative impression was partly due to the apparent lack of effort in his gestures (d’Escrivan 2006), but some in the audience (most of them considerably younger than I) found this knob twiddling pretty cool and expressed their opinion with “Yeah!” and other positive exclamations. I suppose that in a culture where manual manipulations of game controllers, mobile devices, remotes, touch screens and all manner of keyboards are central to daily life, the appreciation of a well turned knob should not be surprising...

The second example from the RIJF was The Dave Rivel.lo Ensemble (Rivello, 2013), a totally acoustic big band that serves as a vehicle for Rivello’s beautiful writing and arranging. One piece featured a solo by Mike Kaupa, a very fine trumpet player. About halfway through the piece I noticed that Kaupa had reached down, picked up a small instrument mic and clipped it to the bell of his horn, which was puzzling, as the other soloists had used stand-mounted mics set up on stage. When he began his solo, it became clear that he was using a loop delay, which is unusual in a big-band setting, despite the fact that the first use of a loop delay in jazz was by trumpeter Don Ellis and his innovative orchestra in the late 1960’s (Ellis, 1967).

Kaupa’s solo differed from the looped solos taken by Panzani in that he had to fit his solo into a denser and more fixed background from the big band than the free background provide by the other 2/3 of THIEFS. Simply put, THIEFS was agile enough to change the background to fit the looped solo, but Kaupa’s trumpet loops had to fit the tempo and tonality of the background that Rivello wrote for the trumpet solo. Rivello’s approach was similar to the approach Ellis used so long ago, which was to make the background static, both tonally and rhythmically. Rivello used a catchy ostinato, while Ellis tended to use sustained chords with no rhythmic pulse. Since Kaupa’s loops were not synchronized with the tempo of the piece, his solo had a polyrhythmic quality that worked well over the ostinato.

So far so good, but toward the end of Kaupa’s solo he stopped looping and switched to a harmonizer. A sense of THIEFS deja vu struck me as Kaupa’s solo came to an abrupt halt while he tapped his foot 11 times (I counted), apparently through a menu, to get to the harmonizer setting. He then started playing again, a bit tentatively at first, but quickly got back into something interesting. It was a repeat of Panzani’s transition from loop delay to harmonizer with THIEFS, albeit a bit less intrusive visually and in a maximally acoustic setting, which made it totally unexpected.

Zero-Sum Game

The bottom line on these two vignettes is that the technology used to extend improvisation had unintended consequences on the improvisation itself. There seems to be something of a zero-sum game (at best) in the use of technology in jazz improvisation, both for the performer and
for the audience. When a performer has to manipulate a
technology interface, the manipulation of the instrument
interface seems to suffer. To be sure, it’s hard enough to
truly improvise well, requiring considerable multi-sensory
concentration (Berliner 1994), without the visual, cognitive
and motor distraction of cycling through a menu with one’s
foot. As baseball great Yogi Berra famously said, “You
can’t think and hit at the same time.” (O’Toole 2011).

Some of this is due to the questionable usability of the
interfaces on the effects boxes. These user interfaces usually
are designed around the functionality the boxes provide
rather than the musical tasks they will be used for, which
makes their manipulation a non-musical activity. Most in-
terfaces also suffer from controls intended for foot ges-
tures, or they use knobs, which require hand gestures.

It is true that many effects boxes will respond to MIDI
signals that can be sequenced and played back during the
tune to schedule parameter changes. However, fixed se-
quences only work for rigidly timed performances and
break down for improvisational pieces that can’t be pre-
dicted beforehand. Real-time interactive environments
from the electroacoustic world like MAX/MSP (Zicarelli
2013) provide enormously powerful tools for creating arbi-
trarily complex and flexible interactive pieces where pa-
rameter tweaks to software effects can be triggered by arbi-
trary performance gestures, but the technical literacy re-
quired is beyond mainstream jazz musicians.

This complexity zero-sum game seems to apply to audi-
ences as well as performers. When effects are used, the
complexity of the instrument content needs to decrease in
some rough proportion to the increasing complexity of the
effects to avoid overwhelming the listener with complexi-
ty. This inverse relationship was demonstrated in both
Panzani’s and Kaupa’s improvisations. For example, when
Panzani used a timbre modulation with a lush reverb, he
slowed down a bit and articulated his notes more simply to
let each note unfold. When he used the loop delay, he
played much simpler figures as sources for his loops, and
the live phrases he then played over the looped background
were simpler than those he played without the loop delay.
Similarly, the phrases he played into the harmonizer tended
to be short, highly rhythmic fragments rather than the long
complex lines he played when using minimal effects.

**Performing with GenJam**

This brings us to my experiences performing with GenJam.
My performances over the years fall into one of three cate-
gories: background (receptions), mid-ground (cafes and
coffeehouses), and foreground (concerts, clubs and de-
mos). In large venues like a ballroom or RIT’s field house
(where we are ear candy for potential students and their
families at recruiting events) I use the house sound system
and a wireless mic/ear-bud setup, which lets me wander
around the venue. In clubs and coffeehouses, I use the
wireless mic so that I can wander, but I often bring my
own sound system, and I may not need the ear buds. In
foreground concert/demo situations I might use a stand-
mounted mic, since I’m likely tethered to a specific loca-
tion. In all cases, I will have a small setup of a laptop,
pitch-to-MIDI converter, tone generator and small boxes
for the wireless devices. All device displays are facing the
audience whenever possible. I only announce tunes in
foreground settings, but at receptions I often chat with in-
terested listeners who start conversations with me, usually
between tunes or while GenJam is taking a full chorus.

When I perform a tune, I type a line of text on the lap-
top, hit return, and usually wander away from my gear as
the tune begins. I seldom look at the gear during a tune,
and I never touch the keyboard or any other control during
a tune. I plan my wanderings to lead me back to the gear
when the tune ends so that I can start the next tune. In
short, there are no obvious performance gestures on my
part during a tune to highlight the interaction.

This presents the audience with visual cues indistin-
guishable from a classic music-minus-one setup, which, for
most audience members, results in a total lack of aware-
ness that the performance is interactive. The problem is
that there are no visible performance gestures (Hook 2012,
Schloss 2003) that correspond to GenJam’s improvisations.
There is considerable interactivity between GenJam and
me, but it is all aural, not visual. My visible performance
gestures are limited to playing the trumpet, which is opti-
mal for me as a performer but makes the interactivity im-
perceptible to all but the most attentive listeners.

When playing receptions I frequently get questions from
curious listeners who look at the gear and ask, “Where’s
the CD?” I usually reply with, “Well, it’s a real-time inter-
active performance system using genetic algorithms to…”,
which makes their eyes glaze over, and they say, “That’s
nice. Where’s the CD?” Clearly, they can’t wrap their
heads around the fact that GenJam listens to me, actively
develops what it hears me play, and truly improvises.

Occasionally, during a chorus of fours, someone will
ask, “How did you know it was going to do that?” which
conveys that the listener realized that GenJam played
something related to what I had just played. I typically re-
ply, “I didn’t. It heard what I played and developed that in-
to its response.” That usually results in a moment of reali-
zation, and I now have a knowledgeable listener.

Ironically, the clumsy performance gestures detailed in
Panzani’s and Kaupa’s solos highlight the fact that they are
interacting with technology, and the blatant nature of the
loop delay and harmony effects boxes make the presence
of technology that much more apparent. Indeed, most audi-
ence members seemed to focus on the presence of the
technology, which they judged as either inherently good or
bad, rather than on the music enabled by the technology. For them, as is the case for some experimental performers, the technology was an end in itself.

**Classification**

Putting all this together leads me to propose a two-dimensional classification scheme for technology performances, which serves as a framework to contrast performer and audience perspectives. The dimensions are spontaneity and interactivity. *Spontaneity* refers to the degree to which the content is generated in the moment. *Interactivity* refers to the degree to which the content is influenced by interactions with a human performer.

![Figure 1. Scatter Plot of Jazz Performances](image)

A Canned MIDI sequence (GenJam rhythm section)  
B Uniform-random pitch & rhythm  
C Human-operated loop delay  
D Human changing harmonizer setting  
E George Lewis’s Voyager (Lewis 2000)  
F GenJam soloing on Giant Steps by itself  
G GenJam soloing on a tune after Al has played  
H IGA GenJam in audience-mediated performance  
J GenJam trading fours  
K GenJam intelligent echo  
1 Mike Cottone Quartet (straight ahead)  
2 Tim Berne’s Snake Oil (intensely far out)  
3 Big band playing Glen Miller chart (no improv)  
4 Jam session playing Bye Bye Blackbird  
5 Acoustic Piano

*Figure 1. Scatter Plot of Jazz Performances*

Figure 1 shows a scatter plot in this 2-D space of several jazz performances from the performer’s perspective, some that use technology (denoted with letters) and the rest that use no technology (denoted with numbers). The non-technology examples serve to place the technology examples in a broader context of jazz performance. Note that the scales for both dimensions are ordinal only, not interval, much less ratio.

The extreme cases help define these dimensions. At the origin of the graph is example A, a canned MIDI sequence, which is neither spontaneous nor interactive and will play the same every time it is triggered. While it’s true that a MIDI sequence can be triggered interactively by a performer or spontaneously by a software process, the sequence itself is canned. Incidentally, the rhythm and head sequences of a GenJam performance fall into this category.

Example 5, an acoustic piano, is highly interactive, but hopefully very predictable, as should be the case for any traditional musical instrument. B is a hypothetical system that triggers pitches, note lengths, note onsets, densities, etc. using a uniform random number generator. Such a system would be maximally spontaneous but not interactive.

Example E, Trombonist George Lewis’s highly innovative Voyager system (Lewis, 2000), is nearly as spontaneous as B, but it is highly interactive and way more musical. Incidentally, GenJam acquitted itself quite well versus Voyager and two other systems in a recent doctoral thesis on the evaluation of creative systems (Jordanous 2013).

Examples C and D are the Panzani and Kaupa effects vignettes describe earlier. 1 and 2 are two non-technology quartet performances from the RIJF. The fact that these performances are adjacent to one another on the graph drives home the fact that the scales are ordinal, because these two performances were at nearly opposite ends of the jazz spectrum aesthetically from an audience perspective.

Example 3 is a performance of a Glen Miller tune where the solos are played “just like the record.” Example 4 is my sitting in with the Bob Sneider trio at an after-hours jam session that occurred every night of the RIJF after the scheduled shows. Playing at a jam session might seem to be freer than playing with an established quartet (1 and 2), but most guests play it safe in unfamiliar surroundings.

This leaves the GenJam examples, F-K. F, G and H are variations of full-chorus GenJam solos. F is Giant Steps from (Biles 2007), where GenJam plays the head and four solo choruses of the classic John Coltrane tune autonomously, hence no interaction.

G is the more typical case where GenJam takes a solo chorus after I’ve played the head and taken a full chorus myself. During the head and my chorus GenJam listens to my playing and modifies its measure population as detailed in the demonstration at this workshop (Biles 2013).

H is an audience-mediated performance (Biles 1995), where the audience serves as a mob mentor, so to speak, using red/green paddles to signal feedback that guides the evolution of “their” soloist in a live-performance setting. Solos from the resulting soloist, then, reflect this interactive training by the audience.

Example J is GenJam trading fours with a human sideman, as detailed in (Biles 2013). Briefly, GenJam listens to the human with a pitch-to-MIDI converter, builds measure and phrase chromosomes for the human’s four, mutates those chromosomes using its musically meaningful mutations, and immediately plays the result as its response in the next four.
Finally, example K is GenJam’s intelligent echo mode, where it in essence performs an intelligent loop delay of the human’s improvisation that maintains the tempo, tonality and eighth-note quantization of the tune. This eliminates the problems with loop delay effects that simply record and play back exactly what the user played between the record start and stop signals, and it even eliminates the need for the human to signal start and stop, as GenJam loops everything the human plays during the echo chorus.

**What about the Audience?**

Figure 1 represents classification from the performer’s perspective, which is relatively straightforward because the performer presumably has a clear understanding of how the technology (if any) works and how the piece is supposed to proceed. The audience, on the other hand, generally doesn’t have that knowledge and is left to interpret the performance using whatever knowledge, experience, expectations, biases, etc. they may bring to the table. This can turn performances into projective tests for many audience members when they are unprepared for challenging or unexplained content. To help get a handle on how mainstream audiences process a performance I’ll borrow from the UI world (Biles 2007).

Don Norman’s model of user interaction (Figure 2) is an elegant conceptualization of how designers and users interact through the system image of a piece of technology. Designers have a mental model of the system they build, and that model includes design goals and constraints, implementation details, intended uses for the system, and other under-the-hood aspects that are not readily apparent in the resulting system and that would likely confuse users if they were required in order to use the system.

![Figure 2. User Interaction Model (Norman 88)](image)

The aspects of the system that the user can experience directly make up the system image. This includes not only the user interface and functional behavior of the system, but also any documentation, marketing material or reviews the user might access. The user’s mental model begins with any knowledge, experience, instincts, or other cognitive content they apply in using the system, and it changes (hopefully for the better) as the user interacts with the system and learns to use it to perform a task. A usable system, then, is one that an intended user can readily manipulate to enhance the performance of some desired task.

Norman’s model can be adapted easily to jazz performance situations as shown in Figure 3. The designer is the performer; the user is the audience; and the system image is the performance itself, including all aural/visual aspects, the venue, the performer’s gestures, program notes, etc.

![Figure 3. UI Model Applied to Jazz Performance](image)

The performer’s mental model informs the creation of a performance, as well as the classifications shown in Figure 1. The audience doesn’t have the performer’s mental model at their disposal, so they must process a performance using whatever knowledge and expectations they have in their initial mental model to interpret what they experience.

This audience mental model is tough to pin down, because different audience members have different mental models, which is part of what makes engaging an audience so challenging and rewarding. To demonstrate how the performance examples in Figure 1 might be classified by the audience, I’ll borrow again from the UI world and propose an audience persona (Cooper 1999) named Fred to provide a hypothetical but realistic audience member to focus on.

Fred is a 46-year-old marketing account manager who loves music and has eclectic musical tastes. He is not a musician, but he listens to a lot of music and considers himself a jazz fan. His favorite jazz musicians include Dave Brubeck, Chris Botti and Norah Jones, all of whom he has enjoyed at the RIJF. He usually buys a club pass for the RIJF and attends most of the evenings. He also usually buys a ticket to one of the headliner shows, this year Bob James & David Sanborn with Steve Gadd. Conveniently, he also caught all the RIJF shows referred to in this paper, and he has seen a YouTube video of George Lewis with Voyager. On a whim he attended a GenJam set at the Rochester Fringe Festival (Fringe 2013). Actually, Fred is an amalgam of dozens of people I’ve met at RIJF and/or GenJam performances over the years. Figure 4 is Fred’s interpretation of the examples presented in Figure 1.

![Figure 4. Scatter Plot from Fred’s Perspective](image)
Fred conflates the inside vs. outside contrast between the non-technology quartets 1 and 2 by seeing 2 as more spontaneous and interactive than they likely are, while misinterpreting the straight-ahead aesthetic of 1 as being less spontaneous and interactive than they are. He credits C and D with a bit more spontaneity than they likely deserve, but he does correctly map the foot tapping and knob twiddling described in the Panzani and Kaupa solos to their aural results. It’s the GenJam examples, though, where Fred shows the greatest divergence from the performer’s perspective.

Fred totally misses the spontaneity and interactivity of the full-chorus GenJam solos (F and G) and thinks that they were simply recorded. He gets the audience-mediated performance example (H), but he totally misses J, GenJam trading fours with me. As described earlier, GenJam trades fours invisibly, and since the interactivity is not obvious, Fred doesn’t perceive the spontaneity either. He assumes that GenJam’s fours are canned, which places them near A and totally undervalues GenJam’s abilities.

Fred does get example K, GenJam’s intelligent echo, because it is the most transparently interactive of GenJam’s improvisation modes, especially when I trade ones with it for a few measures. Fred might be a bit confused, though, when I play a melodic figure followed by the figure transposed up a third, which sounds to him like a harmonizer. When I play counterpoint or just play a lot of notes, Fred doesn’t perceive the looping at all, which might actually make it seem more spontaneous but less interactive.

Fred’s mapping of performance examples onto the spontaneity-interactivity space is one hypothetical example, but other audience members (or readers of this paper) might produce different mappings. To collect data with this scheme, one could present a set of performance examples to subjects, ask them to rank order each performance on the two dimensions separately, and plot the resulting ordinal rankings as in Figures 1 and 4.

Finally, to address my audience perception conundrum with GenJam, there are some things I can do. In foreground settings 1 I can explain how GenJam works between tunes, and in demos I can use a visualizer to animate GenJam’s and my notes as different colored balls drifting across a 2-D display. I often plant Top-10-Things-about-GenJam brochures around the venue, but they typically are ignored. Indeed, mainstream audiences usually see these attempts as academic intrusions on the music. They came to listen!

All this gets at the heart of the problem. By performing in an established genre with well defined audience expectations that typically exclude technology, GenJam’s seamless performer interface renders its interactivity and its spontaneity imperceptible to most mainstream audiences. This used to bother me, but I’ve decided that if they thought GenJam’s playing was recorded, it must sound pretty convincing, so I take this lack of appreciation as a complement. After all, it is about the music, not the technology!