

Balancing Defiance and Cooperation: The Design and Human Critique of a Virtual Free Improviser

Ritwik Banerji

Center for New Music and Audio Technologies (CNMAT)

Department of Music

University of California, Berkeley

ritwikb@berkeley.edu

ABSTRACT

This paper presents the design of a virtual free improviser, known as “Maxine”, built generate creative output in interaction with human musicians by exploiting a pitch detection algorithm’s idiosyncratic interpretation of a relatively noisy and pitchless sonic environment. After an overview of the system’s design and behavior, a summary of improvisers’ critiques of the system are presented, focusing on the issue of balancing between system output which supports and opposes the playing of human musical interactants. System evaluation of this kind is not only useful for further system development, but as an investigation of the implicit ethics of listening and interacting between players in freely improvised musical performance.

1. INTRODUCTION

“...a so-called ‘pitch follower’ — a device known to exercise its own creative options from time to time...”

— George E. Lewis, [1]

Since George Lewis’ *Voyager* [1, 2], researchers in computer music have designed a variety of virtual performers of free improvisation [3, 4, 5, 6, 7, 8, 9, 10, 11, 12], interactive music systems built to perform (ideally) as just another semi-autonomous musician in an ensemble of human improvisers. As Lewis writes regarding *Voyager* [2], virtual improvisers should listen and respond to human playing to produce sonic output which appropriately shifts between supporting and opposing other improvisers’ musical ideas. At one extreme, the improviser should feel the influence of their playing on the system’s output, or as Michael Young describes it, a sense of “intimacy” in the human-machine interaction [12]. At the other extreme, the player should neither feel that the system simply mirrors their behavior, nor the need to “prod the computer during performance,” as Lewis puts it. Similarly, Young describes this as the “opacity”, or the lack of transparency in the relationship of system input and output.

However, building a system to interact with human improvisers with a sense of intimacy and sympathy is a special challenge given the tendency of free improvisers to explore timbrally complex material. As several system designers have noted, free improvisational practice frequently features unpitched and noisy sounds and generally avoids musical structures based in pitch, such as

melody or harmony. Likewise, in order to design systems which listen and respond sympathetically to such playing, many researchers have discarded any design approach in which the system’s real-time analysis of the human player’s sound output is solely pitch-based [5, 6, 7, 8, 9, 10, 12]. In addition to pitch, such systems use a variety of spectral analysis tools to decompose the complexity of common practice in free improvisation into several components, such as noisiness and roughness [10].

Given how improvisers often play, this is a logical direction to follow in building systems to exhibit greater “intimacy” in human-machine interaction. Nevertheless, it overlooks the hidden utility of the pitch detector’s interpretation that pitchless sounds, such as a styrofoam ball scraped against a snare drum head, the hum of an amplifier, or a woodwind multiphonic, have a definite pitch, or more often, are simply a run of pitches. While the pitch detector’s interpretations of these pitchless sounds as “pitched” is technically inaccurate, they may provide a means realizing Young’s ideal of opacity in the system’s input-output transformation, a sense of mystery and individualism which Lewis wittily depicts as the algorithm exercising “its own creative options”.

This paper presents the design of a virtual free improviser, known as “Maxine” [4], built to creatively exploit the pitch detector’s odd interpretations of sonic material in free improvisation, an auditory environment for which it is destined to fail. The system capitalizes on the pitch detector’s simultaneously “intimate” and “opaque” interpretations pitchless sounds in order to produce an overall interactivity which balances between sympathetic and oppositional behaviors. After an overview of system design and its resultant behavior, the paper concludes to focus on the critical evaluation of improvisers who have played with the system as part of an ethnographic study on social and musical interaction in free improvisation.

Ultimately, soliciting practitioners’ feedback on such systems is not simply about improving design, but actually deepening understanding of musical practice on a social-scientific and musicological level. Asking players to critique the behavior of a non-human musician allows them to be far more explicit about what they really expect of other players in socio-musical interaction. Because improvisers wish to show respect for the musical liberty of their peers, it is a social taboo for them to explicitly express such preferences to each other [13]. Conversely, because this system is a mere machine, players feel no risk of infringing upon the musical liberty of others by offering critical feedback to the designer [14]. As a result, testing such systems with improvisers provides a unique opportunity for them to candidly express expectations they hold of their peers’ behavior which they are hesitant to ever make explicit in their routine social interactions.

Copyright: © 2016 Ritwik Banerji. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

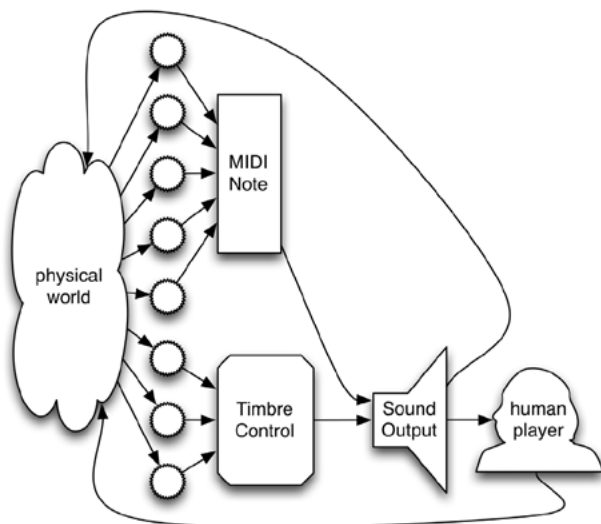


Figure 1. Overall flow of information through system, from physical world, through agents, to sound output and the human performer, and back to the physical world. Small gears represent each agent.

2. SYSTEM DESIGN

The system uses a multi-agent architecture. Several identical agents simultaneously process auditory input and control sonic output. Agents operate non-hierarchically and in parallel to the rest (see Figure 1). While each agent is the same, internal values and end outputs can vary significantly at any given time. Each agent functions as a single “arm” or “finger” of the system, controlling either MIDI note or controller values based on the processing of auditory input from its respective “ear.” The rest of section 2 traces the flow of information through a single agent (see Figure 2).

2.1. Input and Feature Extraction

The system receives audio input from the physical world into two dynamic microphones (Behringer XM8500), one aimed at the human performer and the other at the system’s own loudspeaker output. Analog to digital conversion occurs through a MOTU Ultralite audio interface and audio feature extraction occurs in Max/MSP. Each agent extracts three basic features from incoming digital audio signal in real-time: 1) pitch and 2) attack information, both from Tristan Jehan’s [pitch~] object [16] (based on Puckette’s [fiddle~] [17] and hereafter referred to simply as “[pitch~]”), and 3) amplitude information.

2.2. Note Event Timing Control

Agents use changes in pitch and attack information reported by [pitch~] to control output timing. Reported pitch changes and attacks are sent to a timer, which measures the interval between reported events, regardless of whether the event was a pitch change or an attack. These durations are used to set the agent’s *base quantization*, (BQ). Similar to the *tatum* [18], or temporal atom, BQ is the shortest duration for any MIDI output from the agent. Actual durations, or local quantization (LQ), are a random multiple of the BQ between one and 15. However, not all attacks or reported changes in pitch are used to set the BQ, and reporting of these events is filtered by a probability gate (see section 2.4.).

Commands to change the LQ to a new random multiple of the BQ are sent out at the rate of the current BQ. Another probability gate, however, only allows a percentage of these commands to cause an actual change in the LQ. Similarly, note output messages are being sent out at the rate of the LQ, but another probability gate controls the percentage of these note output messages resulting in an actual MIDI message.

2.3. Pitch Selection

Note output from each agent is selected from a three-value pitch-set within a three octave range (C1 to C4) which may change at any time. A note P_{xc} in the pitch-set $P_{(l, 2, 3)c}$ is changed, to a new value (randomly chosen within same range as above), P_{xn} when an incoming value from [pitch~] P_I is a match (or when $P_I = P_{xi}$). Essentially, this mechanism is much like the modern mechanical arcade game “whack-a-mole”. However, similar to the BQ, not all matches ($P_I = P_{xc}$) trigger a change in the agent’s pitch set because of filtering out by another probability gate.

2.4. Probability Gates

Probability gates determine the likelihood that an incoming message will be passed through the gate. This probability rises and falls according to incoming amplitude from the microphones. However, the mapping is constantly changing both in direction (inverse vs. direct) and in degree (see Figure 3).

Specifically, current incoming volume is scaled to probability according to current high (H) and low thresholds (L) for scaling. Changes in threshold values are triggered by changes in pitch reported by [pitch~]. However, this triggering is also passed through the very same probability gate. Once a change in threshold is triggered, current incoming amplitude data is polled at the rate of the

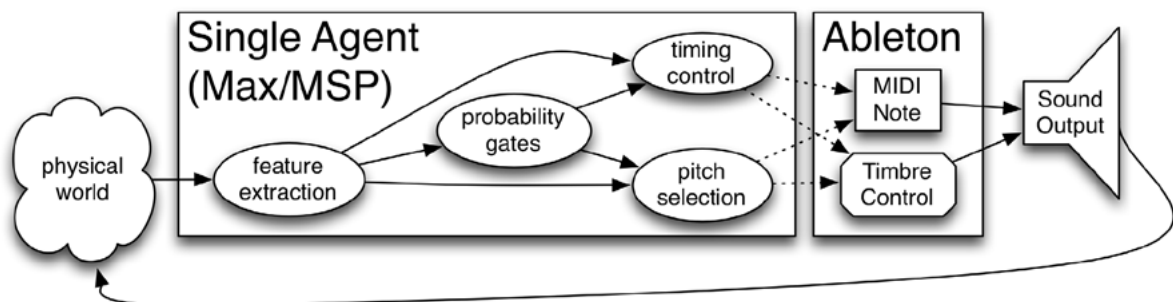


Figure 2. Flow of information through a single agent, from physical world to sound output (and back to the physical world.)

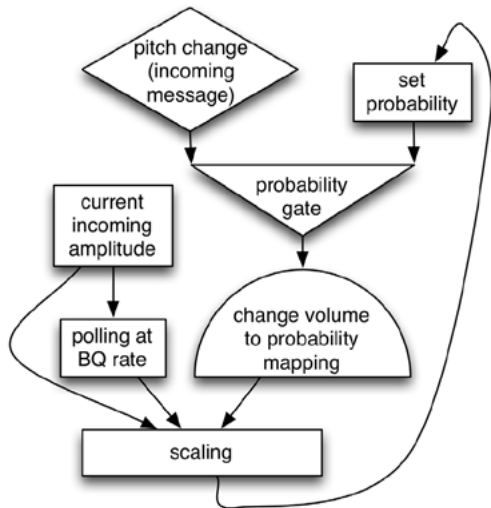


Figure 3. Internal structure of probability gates

current base quantization and sent as the new high or low threshold. When the “high” threshold is lower than the “low” threshold, an inverse mapping of volume to probability results.

2.5. Sound Output and Timbre Control

Based on the above, each agent sends either MIDI-note or -controller values from Max/MSP for output in Ableton Live. In typical performance practice, five agents are responsible for note generation while three control the manipulation of timbre parameters in Ableton. Controller values are used to manipulate the timbre of virtual instruments in Ableton Live. Sound outputs from Ableton Live typically include metal percussion, synthesized versions of prepared or “extended” guitar and piano techniques, a variety of synthesizers, and signal processing tools (e.g. filters, delay, etc.) used to control audio feedback. Strictly speaking, agents controlling timbre only send MIDI notes to Ableton (see Figure 2). These note values are set to control timbre parameters using Ableton’s MIDI-mapping capability.

3. SYSTEM BEHAVIOR

3.1. Intimacy and Opacity from [pitch~]

Because of its reliance on pitch detection as its primary means of real-time analysis of sonic input, the system’s interactive behavior *simultaneously* exhibits both the opacity and intimacy Young idealizes.

3.1.1. Intimacy

[pitch~] offers as a robust, if crude, means of providing this virtual improviser with a real-time analytical representation of its sonic environment. As noted in section 2.2., pitch changes (as reported by [pitch~]) are used to control the relative temporal density of the system’s note output. This allows the system to follow the overall event-density of the human performer (e.g. if the human player’s event durations are between 100-400ms, the system’s output will be in a similar range.)

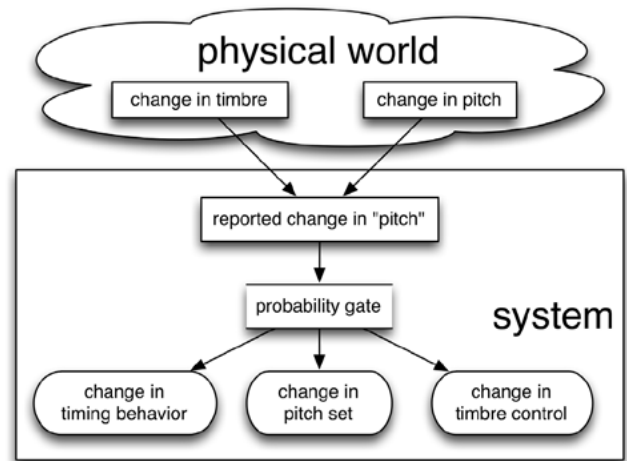


Figure 4. Effect of changes in timbre and pitch in the physical world on system behavior.

3.1.2. Opacity from Intimacy

However, regardless of whether incoming sounds are pitched or not, [pitch~] guesses the “pitch” of these sounds, looking for an even spacing of harmonic partials in order to identify a fundamental frequency. Again, it is senseless to say that a noisy, aperiodic sound like, that of a styrofoam ball scraped against a drum head, has a definite *pitch*. Defiantly, [pitch~] defiantly claims otherwise, reporting an unsettlingly specific value: (hypothetically) a clear D#4, exactly 17.9 cents sharp!

This interpretation is *both* intimate and opaque. It is *intimate* in that for a given spectral profile, [pitch~] will always produce the same estimated pitch value. It is *opaque* in that the relationship of this value and the sound itself, given its pitchless quality, seems almost random.

The system uses [pitch~]’s simultaneously opaque and intimate interpretation of such sounds to produce behavior which is both sympathetic but also mysterious in its interactive logic. For example, given the noisiness of transients at the onset of many note events, [pitch~] often parses such sounds as first a flurry of rapid “pitch” changes (see Figure 4) and then the pitch audibly produced. The human player plays just one note, but the system may respond disproportionately, reacting with a gust of activity to this small perturbation. Still, because of the use of probability gates (section 2.4.), the system does not always react in this unbalanced manner.

For complex, time-varying timbres, [pitch~] allows the system to react temporally to the human player’s subtle modulations of sound. As a player’s timbre changes, [pitch~] yields a new estimated “pitch” value, effectively allowing the system to use [pitch~] as a crude approximation of spectral flux [19, 20], or the degree of variation of spectrum over time. In turn, this enables the system to vary its timbral output and control note activity (section 2.3.), in a manner corresponding to the overall pacing and event-density of the human player.

3.1. Feedback Effects

The combination of the whack-a-mole mechanism for pitch selection (section 2.3.) and feedback (section 2.1.) implies that the system may trigger itself to change its

own pitch set. For example, if the system is producing Gb4 while also waiting for Gb4 to appear in the environment, one would assume that the system has a high likelihood of causing itself to change its pitch set. In practice this rarely occurs. Much of the system's output features heavy manipulations of timbre. As a result, while the system is sending a Gb4 note value to Ableton, the timbre of this Gb4 may be so significantly manipulated that [pitch~] detects no Gb4 from the input signal.

Overall, the use of feedback, in which at least microphone is directed at the system's own output, allows the system to demonstrate a better balance between resistive and cooperative interactivities. In my own early experiments with this system (as a saxophonist), I felt the need to "prod" the system, as Lewis would say, with microphones were only aimed at my instrument. To correct this, the current feedback setup was implemented. This allows the system to respond in a more satisfyingly unpredictable manner as it reacts to its own output, or even the slight hum of the loudspeaker itself. System sounds, alongside other environmental noises, are, like anything picked up by the microphones, just registered as more "pitches" and also stimulate the system to respond. In testing the system with improvisers, microphone placement was often varied based on the player's preference for a more aggressive or sympathetic interactivity.

4. EVALUATION OF THE SYSTEM BY FREE IMPROVISERS

4.1. Methodology

The system was first designed in 2009. Since then, over 90 musicians, primarily in Berlin, San Francisco, and Chicago, have played informal, private improvisation sessions with this system. The initial motivation for testing was to solicit the critique of players with extensive experience in free improvisation in order to identify directions for further development of the system. I approached improvisers directly, as I myself was also a free improviser on the saxophone.

Musicians were asked play a series of duets with the system and give their commentary on its behavior immediately after each piece. These duets were usually between five and ten minutes, though often longer. Pieces typically ended in the same way that most freely improvised pieces do: performers are silent for a period of time and then look up to indicate that the piece is over.

After each piece, I let the improviser lead the conversation, allowing them to focus on whatever they found most interesting or problematic about the system's behavior. As I have discussed elsewhere [14], taking an open-ended ethnographic approach to researcher-subject interaction and letting the subject drive the conversation is far more effective than using pre-determined questions and quantitative evaluation in a controlled laboratory setting. Given the wide range of behaviors desired and resultant from free improvisation (whether human or machine), specific questions and criteria may be irrelevant to the interaction which just transpired and hinder the performer's discussion of their subjective experience of it. As a result, the next section only focuses on performers' comments which directly referred to the issue of balancing support and opposition, and thus is a small cross-section of the

over 300 hours of commentary on this system collected over several years.

This method not only elicited performers' critiques of the system, but their discussion of similar moments of frustration with human players. In other words, asking them to critique the system brought them to express not just what they expect of a *machine*, but what they expect of other *people*. This methodology makes performers feel safe to express such socio-musical expectations in a manner that they never experience in face-to-face interaction with other players. Again, respecting the musical liberty of their peers, improvisers tend to avoid negative critical discussion of their peers' playing. After all, if the practice of free improvisation purports to emancipate musicians from the "rigidity and formalism" [21] of other musical practices, it makes no sense for players verbally express their expectations to other performers, whether beforehand or afterwards.

By stark contrast, players found that critiquing a non-human musician enabled them to articulate expectations that they normally feel implicitly barred from openly expressing in their normal social interactions with other improvisers. While improvisers feel that such direct expressions of expectation are essentially a taboo practice in their socio-musical world, this hardly means that no player has specific expectations, much less that no other player disappoints them. As one performer put it, "I *wish* I could tell other people things like this!"

4.2. Summary of Results

4.2.1. Preference for Greater Assertiveness

On the one hand, many players found the system to be too meek, hesitant, or reserved in its interactive behavior. These performers felt that the system did not take enough initiative in interaction, or as one player put it, failed to "inspire" them. They found themselves stifled by the system's dependence on human input and its tendency, in their experience with it, to wait for the human player to play before producing material of its own.

For example, one player found that the system's silence in some situations was not experienced as a polite gesture of yielding to others, but as a frustrating inability to sustain the drama of the interaction. Such behavior reminded him of an inexperienced improviser whose reticence and lack of confidence saps an improvisation of its overall energy. In response to this sort of system behavior, he stressed the critical importance, in his view, of simply taking a risk and playing *something* rather than remaining silent because of indecision or self-doubt.

For another player, the problem was not that the system was too quiet, but rather that it was *too sensitive* to his playing. Rather than remaining with one sonority for a period of time, the system reacted to his playing too frequently, causing it to change its timbral output too rapidly for this player's tastes. He described the system's behavior as "fickle", even childlike, flitting from idea to idea. While he found moments of the system's behavior interesting, its inability to stay with one sonic idea was disappointing. Instead he would have preferred greater obstinacy in the system's improvisatory behavior, remaining with one idea and allowing the human player to go elsewhere sonically without reacting immediately to each

new idea he introduced. For him, this behavior would have allowed for a more meaningful contrast, juxtaposition of sonorities, or tension to develop.

4.2.2. Preference for Greater Sensitivity

Nevertheless, many other players found the system too aggressive. One individual directly blamed this on the fact that the system reacts to itself. This mechanism and its effects gave him the feeling that the system behaved like a self-absorbed individual during improvisation, following its own ideas rather than cooperating with others. Similarly, another player described this as a failure to “meet me halfway”, or the inability to choose material which partially emulated, and partially deviated, from the choices of the other musical interactant.

In one rather illustrative case, the system persisted with a repetitive undulating feedback effect for nearly two minutes. During this time, the human player experimented with a variety of ideas (melodic runs, sustained tones, quick high-energy blasts, etc.). At one point he stopped playing and stared at the amplifier with a disgusted look, as if to tell the system, “stop!” Indeed, after the piece he described the system’s behavior as “annoying” in its failure to sense his disgust for its playing at that moment.

4.3. Preference for Defiance? Two Individuals

Strikingly, two individuals showed a strong preference for defiant or resistant system behavior. For one Berlin-based cellist, playing with the system was a relatively comfortable experience. Though I asked him if we could do a quick initial piece just to check the volume balance between him and the system, he played with the system without pause for nearly an hour. Such a reaction is unusual, with most players preferring to play with the system for a much shorter period of time, in most cases no longer than twenty minutes.

In the follow-up conversation, he found that he liked playing with the system, but gave some curious reasons for his preference. What he enjoyed most about the interaction was the feeling that the system did not “really *listen*,” as his put it. I was perplexed by this seemingly backhanded compliment. Later on, however, he explained his irritation with players, especially younger musicians, who tend to immediately respond to his playing with material that references (i.e. reproduction or mimicry) what he just played. By contrast, the system’s inability to do so made him feel more comfortable.

The preference for defiant and resistive playing is all the more intriguing in the experience of another Berlin-based musician, this time a trumpet player. As discussed in section 3.4., when a player expresses a desire or interest in more or less aggressive playing, I experiment with varying microphone setups. In three pieces with this trumpet player, each of approximately ten minutes, three configurations were attempted: two microphones on the trumpet player, one on the trumpet player and one on the system, and two on the system.

Surprisingly, he preferred the configuration in which both microphones were aimed at the loudspeaker, away from the bell of his trumpet. Again, this preference, like that of the cellist, suggests a preference for a kind of musical interactivity which is more resistive than supportive. He preferred a manner of playing in which the other

player would be listening closely at all times, but would not necessarily announce and demonstrate their awareness of the other player by immediately and unambiguously reacting to each new idea.

Like the cellist, he explained this preference as a capacity of interaction habit often avoided by younger players. He explained that when he was a younger improviser he too used to play in a much more reactive manner. However, as he became older and more experienced, he reduced this highly reactive tendency in his playing and, similarly, sought to work with players whose modes of interaction with other players were less obvious, or more “opaque,” as Young might describe it [12].

4.4. Discussion

Unfortunately, commentary generated from extensive tests of this system with a variety of players ultimately gives no clear insight into the design of an “ideal” free improviser. Critical evaluations of this system by a wide variety of improvisers reveals a similarly broad range of opinions on how well it balances between engaging in supportive and oppositional behaviors. While the last two individuals discussed in detail above showed a preference for greater defiance and less reactivity in the system, many other individuals asked to critique the system did not agree with this assessment. In the end, data from this study does not indicate conclusively one way or another whether the system should be designed to be more supportive or more aggressive in its interactions with human improvisers.

However, as I have previously argued [14], commentary elicited in tests of this system have a value which goes far beyond simply refining the design of interactive music systems. Asking players to critique the playing of machines built to interact with human players like free improvisers elicits a discussion of what conduct is preferred in these interactions. In other words, the confrontation with a non-human musician brings improvisers to discuss the sense of *ethics* which they enact in how they listen and react (or not) to other performers.

Generally speaking, critical commentary on this system is useful for complicating any simple understanding of behaviors or dispositions such as “sensitivity,” “supportive,” “aggressive,” or “defiant” as descriptors for the behavior of an improviser, whether human or machine. In the case of the improviser who found the system’s behavior fickle, the system’s behavior can be said to be both *too* sensitive and *not* sensitive *enough*. On the one hand, the system’s inability to filter out or ignore what the human player was doing could be described as too sensitive or too responsive. On the other, this unfiltered hyper-reactivity is itself a musical behavior which reflects the system’s failure to interpret the intentions and desires of the other player, a kind of interaction which could also be described as *a lack of sensitivity*.

5. CONCLUSION

The diversity of opinions on this system’s behavior provides no clear answers for how this system ought to be re-designed. Similarly, this broad range of opinions fails to indicate with any finality how one might design an “ideal” improviser that might satisfy all tastes. While such inconclusivity may be frustrating for a designer

simply looking for the best way to build a new system, analysis of the results of this study along these lines is rather short-sighted.

To mine commentary on how this system behaves for insights into further systems' development is to miss the tremendous opportunity it provides to empirically investigate the nature of social interactivity, whether through music or other expressive modes such as language, as specific forms of culture. Specifically, commentary on this system reflects a broad range of *notions of freedom* and ethics which guide how players engage in moment-to-moment decision-making in the course of their improvisatory interactions with other individuals. For those who desired greater sympathy from the system, their opinion reflects a general belief that the autonomy of one individual must be exercised in a manner such that it does not infringe upon the experience of liberty for others. Conversely, those who desired the system to demonstrate greater autonomy implicitly advocated a very different conceptualization of the relationship of freedom and ethics: the more that the system's behavior was uninfluenced and autonomous in relation to theirs, the more they experienced freedom themselves.

Acknowledgments

Field research for this project would not have been possible without financial support from the Mellon Foundation, Fulbright Journalism Fellowship (Germany), the Berlin Program for Advanced German and European Studies. Thanks CNMAT for support for publication of this paper, and to Adrian Freed, Benjamin Brinner and anonymous reviewers for their helpful comments.

6. REFERENCES

- [1] G. E. Lewis, "Interacting with latter-day musical automata," *Contemporary Music Review*, vol. 18, pp. 99-112, 1999.
- [2] G. E. Lewis, "Too Many Notes: Computers, Complexity and Culture in Voyager," *Leonardo Music Journal*, vol. 10, pp. 33-39, 2000.
- [3] G. Assayag and S. Dubnov, "Using Factor Oracles for Machine Improvisation," *Soft Computing*, vol. 8, pp. 604-610, 2004.
- [4] R. Banerji, "Maxine Banerji: The Mutually Beneficial Practices of Youth Development and Interactive Systems Development," *eContact! Journal of the Canadian Electroacoustic Community*, vol. 12, 2010.
- [5] T. Blackwell and M. Young, "Self-organised music," *Organised Sound*, vol. 9, pp. 123-136, 2004.
- [6] O. Bown, "Experiments in modular design for the creative composition of live algorithms," *Computer Music Journal*, vol. 35, pp. 73-85, 2011.
- [7] B. Carey, "Designing for Cumulative Interactivity: The *_derivations* System," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012.
- [8] D. P. Casal and D. Morelli, "Remembering the future: applications of genetic co-evolution in music improvisation," in *Proceedings of the European Conference on Artificial Life*, 2007.
- [9] N. M. Collins, "Towards autonomous agents for live computer music: Realtime machine listening and interactive music systems," Ph.D. Thesis, Faculty of Music, University of Cambridge, 2006.
- [10] W. Hsu, "Using Timbre in a Computer-Based Improvisation System," presented at the *Proceedings of the International Computer Music Conference*, 2005.
- [11] A. Linson, "Investigating the cognitive foundations of collaborative musical free improvisation: Experimental case studies using a novel application of the subsumption architecture," Ph.D. Thesis, Faculty of Mathematics, The Open University, 2014.
- [12] M. Young, "NN music: improvising with a 'living' computer," in *Computer music modeling and retrieval. Sense of sounds*, R. Kronland-Martinet, S. Ystad, and K. Jensen, Eds., ed: Springer, 2008, pp. 337-350.
- [13] D. Borgo, *Sync or swarm: Improvising music in a complex age*: Continuum, 2005.
- [14] R. Banerji, "Maxine's Turing test—a player-program as co-ethnographer of socio-aesthetic interaction in improvised music," in *Proceedings of the Artificial Intelligence and Interactive Digital Entertainment (AIIDE'12) Conference*, 2012.
- [15] T. Blackwell, O. Bown, and M. Young, "Live Algorithms: towards autonomous computer improvisers," in *Computers and Creativity*, ed: Springer, 2012, pp. 147-174.
- [16] T. Jehan and B. Schoner, "An audio-driven perceptually meaningful timbre synthesizer," *Proceedings of the International Computer Music Conference*, Havana, Cuba, 2001.
- [17] M. S. Puckette, T. Apel, and D. Zicarelli, "Real-time audio analysis tools for PD and MSP," in *Proceedings of the International Computer Music Conference*, 1998.
- [18] V. Iyer, J. Bilmes, M. Wright, and D. Wessel, "A novel representation for rhythmic structure," in *Proceedings of the 23rd International Computer Music Conference*, 1997, pp. 97-100.
- [19] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The Timbre Toolbox: Extracting audio descriptors from musical signals," *The Journal of the Acoustical Society of America*, vol. 130, pp. 2902-2916, 2011.
- [20] M. Malt and E. Jourdan, "Zsa. Descriptors: a library for real-time descriptors analysis," in *Proceedings of Sound and Music Computing (SMC 2008)*, Berlin, Germany, 2008.
- [21] D. Bailey, *Improvisation: its nature and practice in music*: Da Capo Press, [1980] 1993.