Ferin Martino: A Small Piano Algorithm and Its Lessons on Creativity, Interaction, and Expression

Jeffrey M. Morris
Texas A&M University
Department of Performance Studies
4240 TAMU
College Station, Texas 77843-4240
morris@tamu.edu

Abstract

The process of developing and making music with Ferin Martino, a piano-playing software algorithm has opened up a number of inquiries surrounding modern performance and creativity. Although it is relatively simple program, it is capable of much unique, expressive output that maintains a consistent style. Instead of constructing musical passages note by note, its musical character arises from its structure, independent of content. Recombination of previously-played material ensures that the first randomly generated notes end up yielding musically and stylistically coherent passages without compromising the variety of musical material produced. A flow of pseudorandom numbers is disturbed by motion in front of a camera, making the music responsive to its surroundings. Human-machine collaboration with this system has been effective in a variety of applications ranging from background music to computer-aided composition and live performance. A brief explanation of the algorithm is followed by explorations of its resulting character; strengths and challenges; and perspectives it inspires on the relationships between humans and machines and on twenty-first century musicianship.

Ferin Martino is a responsive and expressive piano-playing software algorithm with a consistent musical style and playful character resulting from my focus on building structure (rather than content) and allowing environmental motion to disrupt its processes. Originally intended as an installation, human-machine collaborations with it have been effective in computer-aided composition as well as live improvised performance. The algorithm is relatively small: a single-screen Max patch (Favreau et al. 1986). Scholars in algorithmic music are often inclined to ask how musical material/structure and how human motion/pitch input are represented within the system. In short, they are not given any representation beyond their original sensation: streams of MIDI note events and a single video motion value changing over time. The character of the system emerges from simplistic and indifferent processes interfacing with intricacies of organize motion and human engagement. Unlike David Cope's work (Cope 2013), mine is not an expert system gaining its style from statistical tendencies in the music of others; it is instead generative, recombinatory, and influenced coarsely by motion nearby. Its character emerges from its own minimal structure. Unlike David Rokeby's *Very Nervous System* and related works (Cooper 1995), this system does not interpret motion gestures and map them to musical roles.

After a brief explanation of the algorithm, I will discuss the musical character and (pseudo-)personality that emerges from it, the experience of collaborating with the system to make music, and its implications. The algorithm is presented one feature at a time, each along with its results, so that any of these techniques may be applied individually other systems.

Mappings

As with many generative music systems, this one begins with a (pseudo-)random number generator and ends with MIDI output. This random walk is mediated by a flow of data coming from a camera focused on the viewer, collaborator, passersby, or general environment: Frame differencing and pixel averaging through the `cv.jit.mass` external object (Pelletier 2004) yields a single value roughly correlated with the amount of motion in front of the camera. This motion value controls the rate of notes played as well as how long to wait before sampling the video motion again, speed varying directly with motion. Motion is also mapped directly to MIDI velocity (loudness). This is the extent of the "content"-producing portion of the algorithm. From this proto-musical kernel and flow of influence, a number of rigid and non-intelligent processes ensure a coherent structure.

Mode Switching

Each rigid process switches between two modes. **Texture** arises from a randomized decision to play either a single pitch or a chord, each time a new event is triggered. This decision point is slightly weighted toward making single pitches. **Harmony** is dictated by a random walk that selects a new interval each time a chord is triggered. That interval is used to create an equidistant trichord, placed below the most recently played melodic pitch so that together they would create an equidistant tetrachord. It is probably this momentary intervallic coherence that makes the harmonic character reminiscent of the free atonal music of the Second Viennese School.

Another process lays a fertile ground for motives and musical ***sequences*** by regularly switching between playing new notes or letting a stream of previously played notes flow through to the output instead. If this switch stays open to past notes long enough, a feedback loop arises, generating sequences. Each repetition is offset by a small interval (randomly chosen, up to a few semitones in either direction) so that when sequences occur, they will likely build dramatic upward or downward gestures.

Since this switch is not sensitive to musical context, it is unlikely that a motive (as we might identify it) will be recaptured in full. Instead, the middle or tail end of it will bloom into a musical ***fortspinnung***. When the switch stays in this recall mode for very brief moments, it has the musical result of internal repetition within a motive, a kind of musical rumination. Both of these phenomena contribute greatly to a sense of coherence at the scale of a few seconds to half a minute.

Ferin Martino's Character

Expressivenes

There is more to its character besides the musical constructs described above (intervalic unity, two-handed textures, sequences and motivic rumination). Since motion detected by the camera is mapped to tempo, the software often responds in direct ways to movement it sees, as if accompanying a free-form dance. However, since that tempo also dictates the rate at which motion is sampled, the software will continue at a constant rate for brief periods at a time and then often change drastically. This allows capricious rhythm to emerge: stable enough to suggest that there is a pulse organizing the music, and dynamic enough to allow us to perceive wide leaps in tempo as changes in rhythm (e.g., eighth notes to half notes) with slight tempo changes or expressive *rubato*.

Mapping motion to MIDI velocity also allows for loudness curves that are mostly consistent but never precisely constant, and occasionally have dramatic bursts. The former evokes a believable sense of intentional dynamic level, and the latter suggests believable accents. This has led pianist colleagues of mine to observe that this performer likes to play a lot with "color."

Playfulness

Playfulness Emerging from Mode Switching Beyond expressiveness, Ferin Martino is playful. Ironically, this human attribution arises from the rigid, non-intelligent, context-indifferent mode switching between playing new material and reusing old material. Besides making its own pitch decisions, software is able to take input from the user through a keyboard or a microphone, with pitch detection employing the `pitch~` external object (Jehan and Schoner 2001), represented as streams of MIDI note messages over time. This feature combined with the rigid mode switching between old and new material means that sometimes the software will embrace the user's input by repeating it once with some variation and then moving on, sometimes it will take that input and make it bloom into a dramatic sequence, and sometimes it will leave space around the user's input but proceed without echoing or adopting the new material. In most cases, Ferin will capture part of the user's input (because the new/old mode happened to change in the middle of the user's input) and incorporate it into the musical rumination. We typically don't intend for machines to ignore our input. When kept in moderation (it never does this for long) and when presented in a continuum of possible reactions, the result can be perceived as playful and independent, rather than dutiful or dumb.

Playfulness Emerging from Normalization Another aspect of this seeming playfulness resulted from a practical need to periodically normalize the range of motion detected by the camera. If the software did not periodically adjust to the normal range of motion values, it would often be overly active or inactive. This results in two effects that add to a playful character. If left alone, it will grow quiet after a visitor leaves, then the software will eventually reset its normalization. If there is no physical motion in front of the camera, the software will become sensitive to the noise in the video signal and become highly active and dynamic, as if entertaining itself. When another visitor approaches, Ferin will typically display a burst of activity and then stop when it normalizes to the new level of motion, as if excited or startled by the visitor (or perhaps trying to attract his or her attention) and then standing and staring at the visitor as he or she stands and stares back. Gradually, both will settle into a more comfortable relationship and more moderate music as Ferin adjusts to the new viewer's range of motion or the viewer reaches out to interact more actively.

The second aspect of playfulness arising from rigid periodic normalization is that Ferin will eventually appear to become bored with too much consistency in a viewer's motion. If the viewer constantly waves arms around, trying to make Ferin more and more active, it will adjust to this when normalization is triggered and will leave the viewer hanging, or, better conceived, leave him or her to solo as Ferin drops back to provide an accompaniment. If a viewer stands and stares for long, Ferin will appear to bore of that and become more sensitive to small, mostly unintentional, motions of the viewer, for example shifting weight, tilting the head curiously, or scratching an itch.

Playfulness Emerging from Nonlinear Sensitivity to Motion Using frame differencing and pixel averaging from a single perspective is bad at precisely determining the actual amount of motion occurring, and that is wonderful in this case. Imagine holding a plain yellow pencil in front of the camera. Pointing the tip at the camera and moving the whole pencil vertically or horizontally will yield small amounts of detectable motion, because the viewable area of the pencil is small. Point the tip at the ground and vertical motion will yield similarly low motion values since only the pixels near the top and bottom of the pencil are changing (the entire middle section of the pencil looks the same). However, horizontal motion while pointing to the ground yields much more detectable motion as from each frame to the next, background image is replaced by pencil image across the whole length of the pencil. In either orientation, mov-

ing the pencil directly toward the camera would yield very low detectable motion. Natural circular hand motions contain a counterpoint of change in all three dimensions, such that it is hard for us to notice each dimension individually. When Ferin pares away one such dimension and responds to it, then the counterpoint becomes apparent. Counterpoint is a delicate balance between sameness and difference: contrapuntal lines are different enough to seem independent but similar enough to seem like they belong in the same system, as if they are aware of each other like partners cooperating on a task.

Lighting also plays a role. Motion causes highlights and shadows to form, but it is difficult for a viewer to imagine what would be seen from the camera's perspective. Highlights and shadows modulate the detected motion in ways that seem unpredictable but are still tied to the viewer's motion: they form a counterpoint to the viewer's motion instead of a faithful representation of it. I have observed this responsiveness being influenced by hair, skin, and clothing color; the dangling of hair, clothing, or jewelry; presence and motion of bystanders or passersby; or even the weather, as clouds mediate ambient natural light coming in through windows.

Further, the playful feedback loop between the viewer's motion and Ferin's responses evokes an interesting variety of motions in the user. Ironically, I have found it most effective to influence Ferin's behavior by using the hand motions of an expressive choral conductor, even though the software isn't interpreting downbeats, cutoffs, or swells. I believe this is because circular motions are the most sustainable and repeatable, the complexities of motion and lighting allow for interesting organic irregularities, and basic characteristics of conducting motions transfer well between my musical intentions and the character of the motion detected by the software: for example fast/slow or smooth/choppy motions, gradual/sudden starts and stops, and occasional moments of repose or accent amidst relatively consistent motion characteristics. However, whether the user is conducting, waving, dancing, or pausing to look curiously while passing by, it is enlightening to see how this software inspires movement in each sufficiently interested viewer.

Making Music with Ferin Martino

This algorithm has been productive in a variety of music making settings. The following is a brief discussion of its strengths, weaknesses, and special properties in each application.

Installation/Divertimento

This software was original designed to be an interactive art installation to demonstrate our Yamaha Disklavier pianos for facility tours with a modicum of responsiveness to the viewer. The striking character of the algorithm inspired the more formal installation titled, "The Collected Solo Piano Works of Ferin Martino, as Conjured by Your Presence." This title reflects the anthropomorphic character of the algorithm and allows the viewer to ponder the notion of authenticity in creativity in multiple ways, raising questions such as, "If the music I hear is affected by my presence, then what does the music of Ferin Martino *really* sound like—is their a definitive version? If so, I can never hear it." The notions of human and machine roles, the meaning and value of authenticity and creativity, and similar issues are important for twenty-first century artists to consider.

Ferin's character seems to be most effective when there is a reason for people to move near it, for example, playing divertimento music in a light-traffic area or a place where people are performing some other task. I find it to be enjoyable as background music in my office as I work. I occasionally notice when Ferin has latched on to some motion of mine, I notice how that influences the way I move (as I have a fleeting moment of discernible influence over the music), and it makes mundane aspects of my day more artistically mindful. Conversely, Ferin doesn't respond well to the impatient museum exhibit hopper who stares for a few seconds to see if it does any tricks like a zoo animal before moving to the next attraction. When presented as an installation, Ferin's personality is best appreciated when the environment allows it to grow on you gradually over long periods of time. It is more of a playful companion than a show stopper. This has interesting implications for the ways in which we present and observe interactive installation art.

This installation has been exhibited at the Triennale di Milano, Italy (as part of the XVI Generative Art international conference) in 2013 and the 2014 International Computer Music Conference in Athens, Greece.

Computer-Aided Composition

Ferin's ability to generate music independently while imbuing it with organic fluctuations and allowing for user intervention makes it a promising tool for composition. I usually work in complete "takes," improvising with it and either working in my own structured ideas or responding freely to the music that arises.

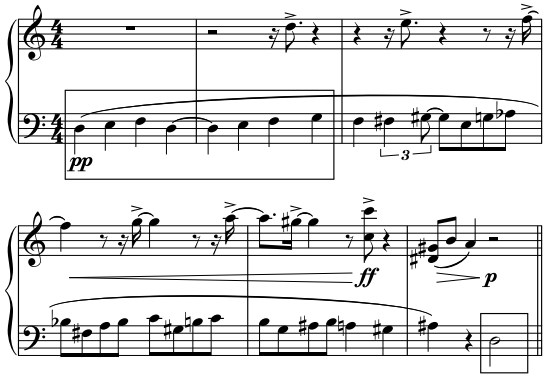

Figure 1: An excerpt from *Canada Fantasia* demonstrating fortspinnung and new material added by the software. Notes within the boxes were performed by me; all other notes were performed in response by the software, all in real time.

In *Canada Fantasia* (2013), I used the keyboard to input motives from the tune, "O Canada" (Lavallée 1880), gradually introducing this minimal content into the musical di-

alogue as opportunities arose. I also "conducted" with my hands to influence the expressive flow as we played together. Figure 1 shows how Ferin responded to my playing the middle portion of the tune (sung to the words, "With glowing hearts we see thee rise"), embellishing and complementing it as well as magnifying the word painting in the passage. It was only by chance that the algorithm caught that motive up in fortspinnung and decided to have the resulting sequence "rise" in pitch as the lyrics suggest.

As it turns out, Ferin's music may sound "pianistic" (in part because of its balance between left hand chords and right hand melody and its strong voice leading resulting from the random walk-controlled harmonic intervals), however it is rather difficult for human pianists to perform. For example, harmonies oscillating among augmented triads, diminished triads, and clusters are difficult to notate in reader-friendly ways, in turn making it difficult for the reader to process and perform the notated music. (Perhaps this phenomenon exposes an unintentional influence of literacy upon piano composition.) Ferin's music is also often dense, harmonically, rhythmically, and texturally, relying on dramatic dynamics to keep textures from becoming muddy. Because of this, I have found it more effective to transcribe Ferin's printed music for multiple instruments. *Canada Fantasia* for example was presented as a piano duo (a reduced score is shown here).

Another work, *The Garden of Forking Paths*, was presented for piano trio. In a way, this allows Ferin's natural inclination to play with "color" on the piano to extend into timbral color through orchestration. In this work, I began with no agenda in mind and responded freely to Ferin's prompts, often complementing or reiterating the most striking motives Ferin had played. Interestingly, it has not proven a simple task to adapt the algorithm to sound natural when composing for other instruments, whether polyphonic or monophonic.

Live Performance

I have also made music onstage with Ferin Martino, joined by saxophonist Jayson Beaster-Jones. In the performance, Beaster-Jones played as he normally would, Ferin controlled a Yamaha Disklavier piano, and I stood in front of the laptop computer running the software, shaping Ferin's performance with my hands, and manually turning on and off Ferin's audio input from the saxophone. We realized that in a duo scenario, a good duo partner will usually play at the same time as the other performer, either taking the foreground, working together to share the foreground, or taking the background. Since Ferin's input was designed for only occasional suggestions of material, the stimulus of a good duo partner was too much. A mute/unmute control is effective for moderating moments of congruence and independence between the two performers.

That is the odd thing about this performance: it was a duo between saxophone and piano, and I was the odd man out, a human onstage that didn't make any sounds. This trio with only two instruments and only two humans, and one "conductor" only focused on one of the other performers ended up being another opportunity for the audience to reflect on what we expect of the composer, the performer, and their tools, things that are worthy of our attention, since they reflect the impact of rapid changes in creativity and technology over the last century.

Work in Progress

Dance

Since Ferin is influenced by motion, we have conducted experiments working with dancers. In this situation, the responsive, expressive feedback loop between the human's movement and the software's music has had an inhibiting effect. The dancers tended to expect longer predictable trajectories from the music so they could plan longer, more coherent or involved motions to accompany the music. If the piano were to fully take the lead, they would be able to remain in deference to the piano, only reacting to it, but the software isn't made to take the lead and keep it for long. Through this, we have started to discover Ferin's deeper nature as an inspiring accompanist, a *collaborative* pianist. We did find it helpful to bring in another (human) musician, able to interpret and predict the dancer's movements and help bridge any gaps in the leading/following dynamic between Ferin and the dancers.

Using dancers also raises practical issues of photography. In contrast to the up-close exhibit viewer or the casual passerby, dancers are more comfortable using a large stage area. This means a greater range of distance from the camera, often changing at a rate faster than Ferin's periodic normalization can keep up with, resulting in extremes of high or low activity as the dancer goes from filling the frame to occupying a small part of the frame, to coming up so close that some of the dancer's body is out of the frame. This inspired us to pan and zoom in order to adjust for these changes and to focus Ferin's sensitivities on certain dancers or parts of dancers at various times—cinematography as a musical control interface. This approach is promising, but it brings its own challenges. For example, panning across a stage or zooming in on parallel lines will be sensed as high amounts of motion in the software, overstimulating Ferin and making it unresponsive. Future experiments in this vein will investigate techniques in controlling colors of materials and backgrounds, controlling lighting and shadow, and infrared lights.

Other Future Work

Ferin's short-term memory to build small-scale structures while allowing musical ruminations to evolve and wander is well-suited to the installation and divertimento scenarios. While coherence across larger time spans is desired in compositions and live performances, a human performing partner can make up for the software's lack of long term memory in a number of ways discussed above. Future work will investigate approaches to enable Ferin to build larger forms on its own. Possible approaches include reproducing the old material/new material mode switch at larger time scales, changing less often, and reproducing longer passages from farther in the past. It may also be useful to implement memory cues, after Lawrence D. "Butch" Morris's *Conduction*

technique for coordinating ensembles of improvising musicians (L. D. Morris 1995; no relation), in which the software or human could mark points in time as they occur and recall that material later, on demand. Butch Morris's work has inspired and influenced many aspects of the inquiries into music making and interaction represented this work.

Future developments will also attempt to adopt the algorithm to compose more idiomatically for different instruments and ensembles without requiring a human arranger.

General Discussion

The algorithm's playful personality, responding capriciously to external input, and its rich variety of output while maintaining stylistic coherence are what moved me to personify the algorithm by giving it an anthropomorphic name. I am not the first to do this; most notably, David Cope attributed both a given name and surname to his algorithm, Emily Howell (Cope 2013). Beyond the notion that every system needs a name, Cope and I appear to feel similarly that our algorithms have some ownership of their output, beyond ourselves in some way. If this is a trend, I believe it indicates a new level of maturity within the Information Age, as we increasingly see ourselves as creators of systems—composing the *composers*—instead of merely using technology to take over undesirable tasks or to do tasks faster than we are able (Morris 2008). As William Seaman puts it, "The artist need no longer seek to define a singular artefact, but instead need develop systems that enable a series of sonic artefacts to become operational and polycombinational" (Seaman 2010, 234). This may be part of a broader development in human creativity, in which popular musicians have been releasing bodies of work under different names (e.g., Uwe Schmidt, discussed in Hofer, 2006), perhaps as a way of processing or insulating their own sense of identity in an age when media creates a market where new composers enter more easily and past composers don't seem to die anymore (Bonds 1996).

Life making music with Ferin Martino has facilitated a number of insights and inquiries by the humans that have encountered or collaborated with it. It encourages us to reflect on the different strengths of humans and machines. Beyond the notion of computers performing more and more delegated tasks sufficiently well, encounters like this give us a chance to remember that, if beauty is in the eye of the beholder, *we are the beholders*—that is our special attribute, on which we are increasingly free to focus. Listeners can now engage so deeply that they *influence* the music, artists steer the software through a sea of possible pathways, choosing one concert—or one body of work—from among a multiverse to bring into reality.

Further, the anthropomorphic qualities emerging from this simplistic mapping and rigid set of routines in a single screen of code remind us that it doesn't take complex systems to please, captivate, or challenge us, because we are simply that good at finding beauty around us. We can release some control, discover how simpler systems can engage our imaginations, and allow the natural voice of the machine to emerge so we can better understand its nature and our changing relationships to machines. I believe these topics are essential to twenty-first century musicianship and reconciling it with our understanding of traditional performance values.

Links

For recordings, scores, and more writings on this algorithm: http://morrismusic.org/ferinmartino

Acknowledgments

Thanks to Jayson Beaster-Jones, saxophone (University of California, Merced) for performing with me in developing live performance applications of this project. Thanks also to the Department of Performance Studies and College of Liberal Arts at Texas A&M University for supporting travel to present these works publicly.

References

Bonds, M. E. 1996. *After Beethoven: Imperatives of Originality in the Symphony*. Harvard University Press.

Cooper, D. 1995. Very Nervous System: Artist David Rokeby adds new meaning to the term interactive. *Wired* 3(3).

Cope, D. 2013. The well-programmed clavier: Style in computer music composition. *XRDS: Crossroads* 19(4):16–20.

Favreau, E.; Fingerhut, M.; Koechlin, O.; Potacsek, P.; Puckette, M.; and Rowe, R. 1986. Software developments for the 4X real-time system. In *Proceedings of the International Computer Music Conference*, 369–373. Ann Arbor, Michigan: MPublishing, University of Michigan Library.

Hofer, S. 2006. I am they: Technological mediation, shifting conceptions of identity and techno music. *Convergence: The International Journal of Research into New Media Technologies* 12(3):307–324.

Jehan, T., and Schoner, B. 2001. An audio-driven perceptually meaningful timbre synthesizer. In *Proceedings of the International Computer Music Conference*, 101–104. Ann Arbor, Michigan: MPublishing, University of Michigan Library.

Lavallée, C. 1880. O Canada [music]. Department of Canadian Heritage.

Morris, L. D. 1995. *Testament: A Conduction Collection* [booklet accompanying sound recording]. New World Records.

Morris, J. 2008. Embracing a mediat[is]ed modernity: An approach to exploring humanity in posthuman music. *Performance Paradigm* 4.

Morris, J. 2013. Canada fantasia. Unpublished.

Pelletier, J.-M. 2004. A shape-based approach to computer vision musical perormance systems. In Lyons, M. J., ed., *Proceedings of the Conference on New interfaces For Musical Expression*, 197–198. National University of Singapore.

Seaman, B. 2010. *Recombinant Poetics: Emergent Meaning as Examined and Explored within a Specific Generative Virtual Environment*. VDM Publishing.