Proc. 43rd International Computer Music Conference, Shanghai: ICMA, pp.351-356. ISBN 978-0-9845274-6-5
Authors' Prepublication Camera-ready copy

Interactive Composition and Performance Framework with Evolutionary Computing

Insook Choi
University of Salford
i.choi@salford.ac.uk

ABSTRACT

Evolutionary models such as genetic algorithms and cellular automata have been well adopted by composers. A challenge still remains in practice, how to convey the dynamics of evolution that can be perceivable through performance realization. This challenge becomes more explicit when adopting an agent model such as swarms in which temporality is implicit in their behavioral patterns governed by self-organizational and social dynamics. In this work a compositional approach and system architecture undertake some paradigmatic shift by situating the evolutionary dynamics at the locus of tone production through an ongoing engagement with a performer. Essentially, swarm agents are given sounds to play with, and performers play with these agents. By articulating a kinesthetic framework in performance, Performance Gestural Articulation Unit is introduced as a generalizable action repertoire with implicit and relative duration. The kinesthetic framework of performance factors is proposed in 6 categories as an abstraction applicable as design requirements for interactive performance. This articulation aspires to AI modeling for human-machine performance with kinesthetic evolution. The paper concludes by summarizing how the abstraction is applied to the composition, Human Voice, as a use case context.

1. INTRODUCTION

Evolutionary computing brings a repertoire of simulations to the general class of complex dynamical systems. With an increase of processing power in modern computing, the horizon of interactive composition and performance opens up with new opportunities given the capacity to handle multiple streams of data from an ensemble of run time simulations. At the same time, experience design demands a meaningful paradigm for audiovisual experience. This demand has an implication in music when we consider the concert paradigm where audiences expect to see musicians playing. Nowadays, novel interfaces for musical expression often take the place of traditional instruments in electroacoustic music. In part, this is an effort to visibly stage performance dynamics, which corroborates evidence that kinesthetic elements in music performance are the basis of emotion, especially for an audience seeing performers interacting with sounding bodies on stage [1]. For those who treasure sound, seeing the effort in the sense of Laban's theory of effort, is perhaps more significant than experiencing media spectacles accompanied by music.

1.1 Hysteresis in Composition and a New Challenge

The term, *hysteresis*, introduced by James Alfred Ewing [2] refers to the observed phenomenon from systems with memory where "the consequences of an input are experienced with a certain lag time, or delay;" for example the element iron when removed from a magnetic field releases magnetization with delay. When applied to a tendency of movement, a current sum of force does not only depend on the sum of the cause, but also on the direction the cause is moving. In a swarm simulation hysteresis is observed yielding *tendency profiles* (see section 4.2) that are used to design agent behavioral rules. Hysteresis has a deep implication in musical tone production.

Shaping tones and musical events by listening is inherently circular with auditory feedback from the kinesthetic interaction between performer and instrument. The choice of state transition is directed through listening to previous states generated by previous actions. The iterative cycle of this circuit builds hysteresis, understanding the previous state, casting it to guide a current state, and planning future states. The chain of kinesthetic interaction produces music, and listening shapes the chaining.

In common practice music, desired hysteresis is composed a priori with what we can refer to as path planning of harmonic progression in a tonal system to build an executable musical structure realized by a performer. The very act of composition is to achieve the desired hysteresis experienced by listeners, encouraging them to build a musical memory and retroactively reshaping musical expectations as time passes. Composers use techniques of repetition and recurrent structures embedded in a formal structure and skillfully instruct performers how to shape and pace musical elements through a notated score. Composers hypothesize the score when executed faithfully will deliver the desired hysteresis for desired musical memory, therefor conveying musical ideas. Ringer and Crossley-Holland [3] define composition as "the act of conceiving a piece of music, the art of creating music, or the finished product. …(M)usical works exist as repeatable entities. …(C)omposition is necessarily distinct from improvisation." This traditional definition reinforces bidirectional tension between composition and improvisation around the question of integrity such that the descriptions of integrity for two ends do not share common ground.

Herbert Brun [4] posed the question in 1970: "Can the composer program musical ideas for a computer, and will the output of the computer have musical ideas?" Now the question comes, for deep relevance of musical ideas in computational design, how to adopt evolutionary computing in music composition with respect to the foundations of performance? This is, in part, the research and creation agenda presented in this paper and in the composition *Human Voice*. Musical ideas are realized by extending a performer's kinesthesia. Human Voice experiments with future requirements for dynamical audiovisual experience scenarios, where the visual component provides behavioral kinetics to accompany a performer and to facilitate the evolving auditory and kinesthetic interaction between performer and agents as a generative mechanism.

1.2 Prior Art

The system and composition presented here are preceded by the prototype system "Wayfaring Swarms" and the composition, *Mutandrum.* Wayfaring Swarms utilizes analog capacitive sensing technology and was implemented in two versions, an interactive performance platform and interactive media platform.

Complex dynamical systems have been favorites for DSP engineers and computer music composers. From modeling an instrument [5] to sound synthesis [6, 7, 8] to compositions [9, 10, 11, 12], the signal features from the range of simple to complex patterns offer an efficient generative capacity. For real time composition and performance, [13] introduces an agent model in a live ensemble presentation where a human performs with robotic agents working through a negotiation process mediated by fuzzy logic and weighted probability distribution. Subsequent versions generate musical materials curated by genetic algorithms and Markov chains.

With an Artificial Life approach, the simulation of distributed agents is implemented in ways the agents cooperatively evolve sound repertoires in computer simulation. In [14] a multi-agent model is presented for musical expressivity: agents diversify simple musical seed material through individual rules, evaluating musical inputs to output preferred expressions. The algorithm simulates listening by agents' musical features. Selection evaluates individuals then applies summation and comparison.

2. SWARMS AND THE COMPOSITIONAL APPROACH IN *HUMAN VOICE*

2.1 The Swarm Model

The swarm intelligence research community studies decentralized yet collective behaviors shown in nature and abstracts a set of rules of behaviors and algorithms, then applies to simulations such as flocking birds [15], robotic systems [16], social cognitive dynamics [17], and optimization of computing processes [18].

The most notable signatures of swarm behavior are self-organization and emergent patterns. The swarm algorithm applied in Human Voice simulates heterogeneous agents' swarming behaviors. The collective behaviors are complex with a wide range of dynamic patterns. The algorithm was developed by Sayama to simulate chemical reactive pattern formation [19]. In Sayama's model, behavioral rules are parameterized and the values for behavioral ranges can be input into a matrix called a "recipe". A heterogeneous swarm can be designed by combining multiple parameter sets for multiple subgroups of agents. To initialize a swarm, typically 100 to 300 agents representing 1 to 5 behavioral groups, are thrown together with constrained random positions and initial velocities. From this initial condition, each agent's motion is dynamically influenced at each time step as it detects other agents within its perceptual range. Agents outside of any perceptual engagement stray in quasi-random motion.

The whimsical behaviors of agents spontaneously bond with others then stray, creating a constant flux of the visual mix, a dynamic field for emergent patterns. While the evolutionary trajectory of a given type of swarm can be anticipated as to its patterns over time, the detailed trajectories and patterns of heterogeneous swarms are not predictable. Often the dominant visual features are clusters of agents forming patterns that mix agents of multiple types. Cluster formation and convergence are phase transitions that represent swarm structural dynamics. Phase transition dynamics are globally characteristic though locally unpredictable. For this reason we applied cluster formation data to generate musical structure at a middle ground level—applying a compositional analogy to Schenker's identification of underlying harmonic structure [20]. Feature data from cluster formation is applied to sound design (see sections 2.3 and 2.4), tracking clusters' size, area, and states of bifurcation and convergence. Table 1 enumerates duration profiles for extracting agent data to detect swarm features and phase transitions. Phase transition dynamics can be manipulated by a performer's perturbations of swarm agents, either to accelerate or to delay transitions to extend the durations of feature formation events. The composition Human Voice applies this technique to realize the Schenkerian concept of prolongation of musical structure.

	Agent Data of emerging swarm features	**Duration Range needed to extract feature data**
1	Number of Agents in each Cluster	15-20ms
2	Energy of Agents	25-35ms
3	Area of each cluster	50-70ms
4	Position of each Cluster in performance space	75-100ms
5	Velocity of each Cluster	80-120ms
6	Cluster Deformation	250ms to 2 sec
7	Cluster Divide	1sec to 3sec
8	Cluster Merge	500ms to 2sec

Table 1: Agent data duration profiles for time ranges required to extract data and to recognize features.

2.2 System Architecture and Components

The architecture consists of three subsystems, each with two or more components as shown in Figure 1: 1) Performance Interaction subsystem consists of a touch sensitive screen, GUI, and swarm visualization with display; 2) Simulation subsystem consists of the swarm simulation and species selection functionality; 3) Sound Generation subsystem consists of feature analysis for swarm formation, Sound Design Patterns to process data from swarms, and sound synthesis and audio display systems. This architecture articulates 1) swarm visualization as a part of performance interaction and 2) feature analysis as a part of sound production. The visualized swarms are displayed on the touch screen through which a performer interacts with swarms. Swarm agents perceive performers' touch points as other agents and respond by adjusting their acceleration. A mixture of colored agents represents the simulation's current state, which provides an important baseline to investigate sound production with swarm dynamics in tandem with visual presentation. Agent data is interpreted to generate control data transmitted to sound synthesis. To control multiple levels of temporal dynamics, pattern recognition and feature extraction from swarm behaviors are incorporated into the Sound Generation subsystem.

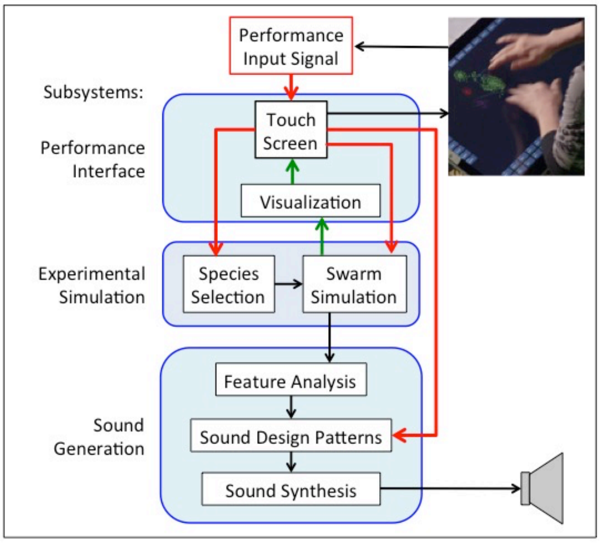

Figure 1: System Architecture. A performer 1) influences agent behavior by manipulation via the visualization, 2) selects agent species bred for the composition, 3) modifies sound synthesis via SDP parameters.

The visualized swarms are projected on a large screen for the audience to see the performance action. The swarms do not visualize the music; rather they are part of kinesthetic engagement for tone production (see section 3).

2.3 Sound Design Patterns

A *Sound Design Pattern (SDP)* is a data-driven building block and a computational model with procedural declaration to accommodate emerging patterns when composing with evolutionary computing [21]. SDPs are designed to work with agent data; they provide the basis for indirection presented in section 4.1. In Human Voice the SDPs primary sound source is a recording of a poet's spoken performance. In distinction to other approaches to evolutionary models, sound microstructure in SDPs is not derived from statistical or other data elicited from the bottom up. Here, the data is elicited from top-down evaluation of agents' emerging patterns on the fly during performer's engagement, then passed to SDPs. This choice of signal flow in performance architecture reflects a hypothesis for collaborative action between swarm agents and human performer towards realizing the musical ideas.

SDP's sound parameter control applies at multiple duration ranges to enable agent data to transform select features and patterns. Duration range refers to signal processing time combined with perceptual time required for a performer to register the attribute or pattern. Figure 2 identifies duration ranges for five sets of SDP attributes. SDP 1: pitch and loudness change, SDP 2: timbre and resonance, SDP 3: sound source location and spatial cues, SDP 4: distinct sound events, SDP 5: patterns of rhythm, tempo, spoken word, and melody. SDP durations are short enough to apply as combinatorial building blocks and long enough to hold coherent musical units. In Human Voice the basic SDP unit is the single spoken word.

2.4 The Compositional dimension in Human Voice

Formal structure associates types of swarms and sounds. Agents' behaviors are studied for temporal signature and classified by tempo (degree of slow to fast) and temporal agility (range of tempo deviation in the stability of swarm dynamics). Temporal signatures are investigated by applying alternating sound synthesis methods, curating musical elements represented as coupled sets of parameters by mapping swarm data to SDPs. Thematic development often relies on recognizable phrase segments with spoken texts functioning as semantic themes. Meanings are not solely dependent on individual texts but emerge from combinatorial bindings according to swarm dynamics. Musical progression is shifted to techniques of progressive derivation of poetic form and linguistic content, enhanced and formalized by swarms' autonomous tendencies. The performer directs swarms to explore boundaries of, and cross modality of, perception emerging from syntactic and semantic properties of language and rhythms.

With the inherent tendency of swarm agents, Human Voice mobilizes emergent discourse by antiphonic and polyphonic responses of multiple voices from recorded words and phrases. The clarity and intelligibility of texts is an important criterion in this composition; extensive voice processing is avoided and timbre transformation of voice is used to help distinguish multiple voices and to enhance the clarity of spoken texts. The choice of onsets and durations of sound sources is based on the temporal dynamics profiles of the spoken texts at phrase level. At all times the number of sound sources are subjected to the continuous tracking of four distinct levels of cluster bifurcation along with swarm data flow. In this aspect, in Human Voice, the dynamics of bifurcation patterns has a structural role in the musical progression while a performer may intervene to prolong, accelerate, avoid or induce phase transitions.

3. KINESTHETIC FRAMEWORK IN PERFORMANCE

Modality that is both auditory and kinesthetic is the most important performance dimension in music. Presenting novel systems is often difficult especially when lacking traditions to bestow performers' and audience's expectations. One of the unsolved problems that linger in working with evolutionary systems is to hear (and possibly see) the emergent properties of the systems through the chain of interaction between human and simulated agents. It is challenging to identify cases where the evolution is bidirectional on the fly between evolutionary models and creative processes, despite facts and cases where arts and music practitioners harvest emergent patterns from complex dynamical systems applied to rich palettes of colors, sounds and structures. It is a hard problem to simulate, or even to articulate, how the evolving dynamics in a performance in situ actually influence their self-referential function during the performance event among all participating agents, with a cascading self-referential framework from individuals to a performance collective.

When introducing a novel interface, in order to seek a solution space we can consult what constitutes a foundation of performance, then translate the problem into system design requirements. A kinesthetic framework can be stated as a set of the design requirements as follows.

1. What moves a performer's state transition? – **Motivational Force** initiates then motivates performer's ongoing movements; a set of choices how to present path planning schema, interface navigation, and sensory motor feedback to performers.
2. What classes of instruments can be referred to with respect to a novel model? – **Instrument Paradigm** aids conceptual orientation for performers as well as the overall design direction for interaction designer. These references serve to frame an initial design space then to break free from it.
3. What areas of a performer's body are in use? – **Anatomical Paradigm** defines characteristics of the human body to specify sensors and actuators.
4. What do performers do with their bodies? – **Action Paradigm** defines gestural constraints inherited from the above items 1-3; a set of specifications for gestural repertoires.
5. What do performers hear? – **Sound Tangibility** suggests kinesthetic responsiveness of sounds; a set of specifications for sound production mechanisms and synthesis methods to engineer a tangible presentation of sounds as auditory feedback.
6. What is the guiding principle from beginning to end during a finite performance event? – **Action Planning Schema** [22], a set of cognitive engineering requirements to facilitate performers to judge their orientation for their next movements. This comes full circle to the first requirements to work with the set of specifications with an added dimension of a full cognitive cycle.

Most of items in this framework are self-explanatory but item 5, and item 6 in relation to item 1 require elaboration. Sound tangibility means more than sound itself. It must convey some kinesthetic affordances suggestive in sounds such that sounds' responsiveness is learnable by performers. This learnability can be facilitated either by consistent sound design techniques or patterns, or by familiar instrumental modeling with physically based techniques, or by other coherent means to achieve rehearsal ability.

Requirement number 6 comes back to item 1 in a full cognitive circle (see section 4 and Figure 2) with an added dimension of understanding and evaluating conditions and criteria, whether performed by computer agents or human performers. The cognitive engineering in this context is comprehensive temporal governance over the system involving all process time from computation to display systems with respect to the temporal reference in a performer's cognitive cycle. Item 6 involves everything in item 1 as performers consult path-planning schema for global guidance, but is primarily occupied, locally refined and tuned, with performers' constructions of their own internal representations of Action Planning Schema. This is acquired by evaluating conditions and feedback to proceed with action selection with projection, meaning anticipation, prediction, and deliberation of future states. With this full circle, performers either experiment with action selection or confidently execute action selection.

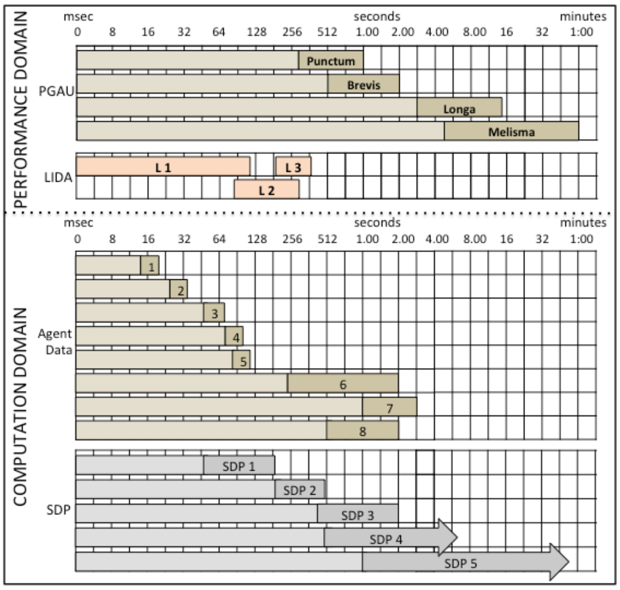

Figure 2: Comparison of duration ranges in performance articulation (PGAU), cognitive cycle (LIDA), swarm agent emerging features (see Table 1), and SDP (see section 2.3).

4. PERFORMANCE ARCHITECTURE WITH COGINITIVE CYCLE

As a context for understanding performers' parallel engagement with multiple temporal layers we refer to the LIDA cognitive cycle of 260-390ms [22]. LIDA identifies a 200-280ms unconscious phase of 80-100ms Perception then 100-200ms Understanding, followed by a 60-110ms conscious Action Selection phase. The exponential timeline in Figure 2 compares LIDA phases with duration ranges of 5 SDPs enumerated in 2.3, agent data from Table 1, and performance gestures (PGAU: see 4.1). A performer's cognitive cycle builds upon short durations of sensorimotor engagements that do not require deliberation, followed by longer durations involving action planning and recognition of musical patterns. SDP1 events occur in a perception time window of 50-100ms, within LIDA's 80-100ms Perception phase. SDP2 events occur from 200-500ms, passing from LIDA's Understanding phase into conscious action. In sum, a 50-280ms cognitive cycle window accounts for perception of pitch, loudness, pulse, and transformations to ongoing sounds.

4.1 PGAU and Model Based Indirection

Swarms are silent. Unlike music instruments' interaction in which performers input excitatory energy into a resonating body, swarms present a system of interaction where a performer can intuit the relationship between the input excitatory patterns and multisensory outputs. The system of interaction consists of performance data, agent data, and associated SDPs yielding an extended interactive pathway through model-based indirection, which applies performance data to efficient transformation of sounds. While learning to play music instruments entails learning to shape the excitatory input patterns, learning to play swarms entails learning the model of inducing changes by steering the social patterns.

Agents do not respond to physical force; they are social. Performance with swarms on a touch surface is experienced more like playing with agents through a fluid transducer as a looking glass entrance to their world. Hand movements on a touch screen are categorized by a unit duration that we refer to as a *Performance Gestural Articulation Unit* (PGAU). In order of increasing duration and indirection, the units are named after Gregorian notation, *punctum, brevis, longa,* and *melisma*. These names are adopted in reference to chant tradition where duration is context dependent. Applicable to both human movements and computational models, PGAU can be defined as *a context dependent unit of gesture with a recognizable pattern having onset and termination in varying duration within a limit, applied to express a functional contribution to temporal dynamics in situ of performance with constituents.* Constituents in this context are swarm agents, SDP, and a performance score.

PGAU is conceived as an expression unit of temporally defined gestural input analogous to excitatory input of a music instrument. PGAU creates a signal that propagates through dataflow in the system, extending its duration in data extraction while influencing temporal properties of SDPs. This way, PGAU sets longer sound events than its own unit duration. Figure 2 implies how indirection works by comparing the relative durations among Agent Data, PGAU and SDP. The greater the duration of PGAU, the more asynchronously the PGAU appears to influence SDP. Model based indirection refers to a signal pathway extending PGAU influence by leveraging the built-in hysteresis of a generative mechanism. Here the generative mechanism is both a swarm model and SDPs.

4.2 Cognitive Cycle in Performance

During the performance with audiovisual continuity driven by swarm dynamics, a performer experiences a time window for action selection. SDPs offer degrees of freedom for PGAU to act in nested time windows with respect to the durations in Table 1 and Figure 2. Prolonged sound responses to PGAUs result from the designed indirection. Swarm state will be extended in sound during the prolonged SDP. The duration of each indirection layer depends on SDP parameters and swarm *tendency profile*; the rate a swarm recovers its native tendency following a PGAU intervention. PGAU onset and musical purpose are the two main action responses for performers to make at any given moment. For action selection, performers require more time to evaluate cluster behaviors and sounds to judge what kind of PGAUs to select, when to onset and where with respect to current bifurcation modes. Locally forming action schema is refined as details of sounds unfold and is cognizant of path-planning schema towards the composition. While a novice can play with this system, performance skill deepens in part through mastering the relationship between PGAU, agents, and SDP through model based indirection. Visual feedback of swarm dynamics indicates states that will influence sounds into the future, and yields an important function for predictive time window for action planning. A performer can estimate the latency from the onsets of PGAUs to the response times in the sounds, also from the releases of PGAUs to the latent duration of lasting effects in the sounds. Figure 3 illustrates a performer's ongoing sensory engagement with latencies from PGAU to swarm states and sounds, while anticipating and planning events in reference to the performance score.

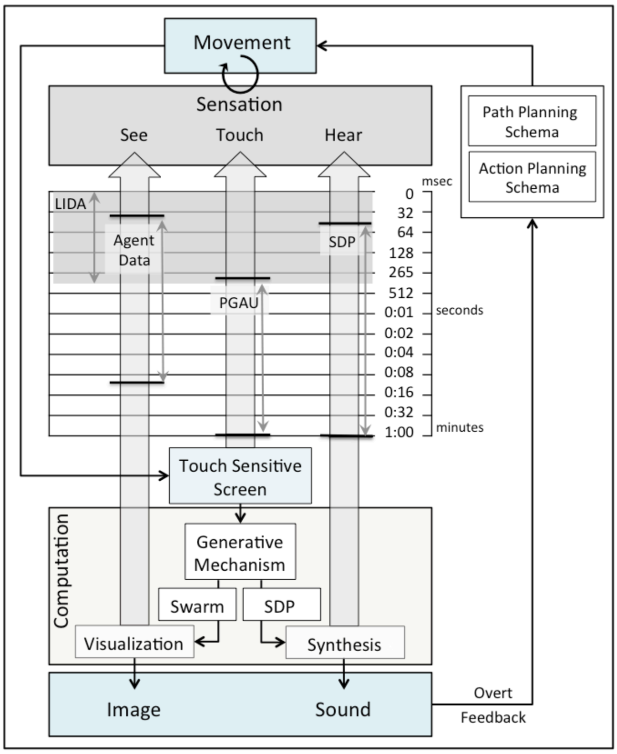

Figure 3: Swarm performance with temporal unit ranges of Agent Data, PGAU, and SDP with respect to the LIDA cognitive cycle. Computational processes sustain the performer's ongoing engagement through sensorimotor feedback (pre-conscious) and overt feedback (conscious). The diagram illustrates a performer's cognitive time windows to process and act upon multiple streams of information.

5. SUMMARY AND FUTURE DIRECTION

The performance architecture is implemented with tripartite requirements: 1) cognitive demand in performance; 2) compositional design with evolutionary affordances; and 3) computational demand for processing evolutionary dynamics for both agent data and SDP. Model based indirection leverages the inherent hysteresis of the swarm dynamics, and the built-in hysteresis of SDPs. This compositional design yields a pathway to extend PGAU signals' impact on resulting sounds with temporal plasticity.

The use case of kinesthetic framework in Human Voice is summarized as follows: 1) the Path Planning Schema in the performance score functions as an initial **Motivational Force** for a performer to engage swarm agents in motion; from thereon, agents provide sustained motivational force. 2) Touch screen interaction with agents has an excitatory quality reminiscent of a plucked and bendable string **Instrumental Paradigm** due to the perceived plasticity. 3) For **Anatomical Paradigm**, a performer's hands provide ten touch-point capacities. 4) To shape excitatory patterns in the swarm world an **Action Paradigm** utilizes PGAU unit duration and intervention repertoire. 5) **Sound Tangibility** is modeled on agent's behavior and coded in SDPs: the ensemble interplay of sound sources is governed by evolutionary dynamics. 6) A performer's sensorimotor engagement with agents' inherent Motivational Force completes the critical round trip coming full circle to **Action Planning Schema** within the larger context of Path Planning, partly evolving and partly in the performance score.

The performer in Human Voice is essentially an agent who brings exogenous energy into the world of swarms. Like machine listening, modeling the external agent's performativity requires defining a universe of principles to shape that energy and how the two worlds work towards musical ideas, which may also evolve. This research is the initial step towards *machine performing*, not unlike machine listening but with cognitive and kinesthetic factors built in the computational automation. This implication extends intelligent instrument design and musical machine learning [23] to model performance disposition where kinetic evolutionary systems exhibit kinesthesia. Therefor understanding a performer's kinesthetic and cognitive factors is a high priority at this stage of experimentation. Evolutionary multi-agents as a generative mechanism provide a context of inherent kinetics applied towards machine performing with human performers. The future direction will 1) continue to adapt evolutionary multi-agents as a generative mechanism, and 2) unfold in formative ways towards realizing machine performing with experimental compositions such as Human Voice.

Acknowledgement

The author would like to thank Sever Tipei for valuable discussions.

6. REFERENCES

[1] Choi, I. "Interactivity vs. Control: Human-Machine performance basis of emotion," *Kansei, the Technology of Emotion.* Camurri, A. ed., 1997, pp. 24-35.

[2] http://www.investopedia.com/terms/h/hysteresis.asp

[3] Ringer, A. and Crossley-Holland, P. "Musical Composition," *Encyclopaedia Britannica,* 2008.

[4] Brün, H. "From Musical Ideas to Computers and Back," Lincoln, H. ed., *The Computer and Music,* Ithaca: Cornell University Press, 1970, pp. 23-36.

[5] Rodet, X. "Nonlinear Oscillator Models of Musical Instrument Excitation." *Proc. of Interntnl. Computer Music Conference*. San Jose, 1992, pp. 412-413.

[6] Chareyron, J. "Digital synthesis of self-modifying waveforms by means of linear automata," *Computer Music Journal*, vol. 14, no. 4, 1990, pp. 25-4.

[7] Kreger, T. "Cellular Automata in Spectral Domain," *Proceedings of Interface 1997*. Auckland, 1997.

[8] Choi, I. "Interactive exploration of a chaotic oscillator for generating musical signals in real-time concert performance." *Journal of the Franklin Institute*, vol. 331B, no. 6, 1994, pp. 785-818.

[9] Beyls, P. "Musical Morphologies from Self-organising Systems," *Journal of New Music Research*, vol. 19, no. 2-3, 1990, pp. 205-218.

[10] Bilotta, E. and Pantano, P. "Artificial Life Models for Musical Applications," *ECAL 2001 Workshop*.

[11] Scaletti, C. *sunSurgeAutomata*. *CDCM Computer Music Series Vol. 3*, compact disc, Centaur, 1989.

[12] Choi, I. *The frog in a machine. Electro Acoustic Music VI*, compact disc, Neuma Records, 1998.

[13] Eigenfeldt, A. and Pasquier, P. "Creative Agents, Curatorial Agents, and Human-Agent Interaction in *Coming Together,*" *Proceedings of 9th Sound and Music Computing Conference*, 2012, pp. 181-186.

[14] Miranda, E. R., Kirke, A. and Zhang, Q. "Artificial Evolution of Expressive Performance of Music: An Imitative Multi-Agent Systems Approach," *Computer Music Journal* vol. 34, no. 1, 2010, pp. 80-96.

[15] Reynolds, C. "Flocks, herds and schools: A distributed behavioral model," *Proceedings of the 14th annual conference on computer graphics and interactive techniques.* ACM, 1987, pp. 25-34.

[16] Martinoli, A. Easton, K. and Agassounon, W. "Modeling swarm robotic systems: A case study in collaborative distributed manipulation," *The International Journal of Robotics Research,* vol. 23, no. 4-5), 2004, pp. 415-436.

[17] Kennedy, J. "Particle swarms: optimization based on sociocognition," de Castro, L. and Von Zuben, F., eds. *Recent Developments in Biologically Inspired Computing*. The Idea Group, Inc., 2004.

[18] Dorigo, M., Di Caro, G. and Gambardella, L. "Ant algorithms for discrete optimization." *Artificial life* 5, no. 2, pp.137-172, 1999.

[19] Sayama, H. "Swarm Chemistry," *Artificial Life,* vol. 15, no. 1, pp. 105-114, 2009.

[20] Schenker, H., *Free Composition*, Oster, E. trans., Hillsdale: Pendragon Press, 1977.

[21] Choi, I. and Bargar, R. "Sounds Shadowing Agents Generating Audible Features from Emergent Behaviors," *Proc. Fourteenth International Conference on Synthesis and Simulation of Living Systems*, MIT Press, 2014, pp. 726-733.

[22] Madl, T., Baars, B.J., Franklin, S. "The Timing of the Cognitive Cycle." *PLoS ONE,* vol. 6, no. 4, 2011.

[23] Rowe, R. *Interactive Music Systems*. Cambridge: MIT Press, 1993