Structured Reciprocity for Musical Performance with Swarm Agents as a Generative Mechanism

Insook Choi[0000-0002-1006-7765]

University of Salford, Manchester, M5 4WT, UK
insook@insookchoi.com

Abstract. A creative inquiry introduces a swarm simulation as a generative mechanism, providing a third agency in music creation alongside agency of composition and performance. When applying an evolutionary model such as swarms to music the challenge is to develop a performance interaction model beyond improvised explorations so that the model can facilitate the integrities for, and emerging dynamics in, all agencies. To minimize a performer's cognitive overload another challenge is management of dataflow in interactive architecture for generating music. An interaction model for structured reciprocity is investigated as a design solution that applies model-based indirection. In the scope of musical composition, reciprocity is exhibited as a time dependent relationship between a performer's actions and evolutionary swarm tendencies applied to generate musical outcomes. Reciprocity is structured using levels of indirection as a construct of mapping data from a swarm simulation to an interactive music performance application. The paper presents an approach to encoding, activating, generating and measuring reciprocity. A comparative case study demonstrates the implementation of these concepts in two musical works. The paper concludes 1) reciprocity in interactive applications helps optimize actions to leverage emergent tendencies, including their intuitive qualities, towards realizing desired outcomes; and 2) data analysis from the case study indicates temporal signatures of reciprocity are related to indirection in interaction scenarios for generative musical performance.

Keywords: reciprocity, swarm simulation, musical performance, cognitive cycle, sound design pattern, gesture, emergent behavior, interaction design, evolutionary model, assistive interface

1   Introduction

Complexity in generative factors in musical performance steers a research agenda for designing interactive experiences. This study aims at gaining insights towards a generalizable principle of structured reciprocity. The principle here is applied to designing efficient and intuitive musical interactions and performance scenarios. Reciprocity in an interactive scenario characterizes a relationship of time-coupled decisions exchanged between a human participant and computation. Cases of creative process focus on where and when decision-making takes place and the transparency of it. A simple interactive system is chosen to emphasize facile real-time selections. The hypothesis is that reciprocity may provide a signature of dynamic selection flow: a process of deferring and taking decisions to render shared creative output.

To investigate reciprocity this work identifies musical composition and performance applications that use generative properties of simulated swarm agents. While musical performance is highly specialized, the use of dynamic simulation in musical creation incorporates multiple constructs also applied widely in interactive media and digital games. To ascertain suitable assumptions concerning musical creation, this case study applies autoethnographic methods to consult the composer and to articulate her criteria for composing musical performance experiences.

The relationship between generative arts and generative mechanisms is commonly understood with the definitions anchored on computational autonomy and causality [1, 2]. Implied in [3] a generative formalism yields how a mechanism can be materialized in arts. Romero et al. [4] preface two different criteria between an artistic perspective and a scientific standpoint. For the former, the challenge has to do with accepting an evolutionary approach through well-established artistic practices and venues. For the latter, the challenge is to innovate and develop autonomous systems that would bring evolvable aesthetics and new models of human-machine interaction.

This paper examines reciprocity as a structured relationship between a performer and a dynamic simulation, where the simulation is applied as a generative mechanism for the interactive production of musical outputs. In musical creation to engage a complex system such as a swarm simulation presents an overarching question: What is the model of engagement? Musical creation draws upon a literature of musical practice, and evolutionary simulation draws upon a literature of computational practice. If the motive for using a simulation is to automate music generation, the automata can be designed with layers of rules and algorithms that influence musical outputs. However when situating a human performer in an engagement model, the design problem space becomes much larger to draw one clean model to work with. For musical performance using generative agents the composer devises ways for performers to engage literatures of both music and computation. Reciprocity is a signature of this dynamic engagement and may be structured to optimize sound production and performance to leverage the

integrity of the evolutionary system in terms of inherent qualities and emergent patterns.

1.1   In and Out of a Musical Paradigm

To study in context a creative technology application such as game design for gameplay experience, assumptions posited will consult criteria and literature of game practice. For computational music creation and performance this paper adopts the same approach. Composition and performance are original sources that bring about interactive sound production and inform listening experience. The fundamental tools of interactive music generation applied in this work are consistent with digital music technology standards. However the works reported here are experimental compositions, meaning they do not reproduce established musical genres but rather develop new classes of references and computational applications. The paper discusses two musical works that present a human performer controlling a "super agent" to interact with autonomous agents in a dynamic visualization of a swarm simulation. In these compositions, swarm agents are given sounds to play with and the performer plays with the swarms to influence the sounds. Agents' collective data controls digital transformations of sounds. Swarms' emergent properties greatly influence the sounds, while the performer can influence but cannot completely control the agents' behaviors. This configuration introduces indirection as a condition for reciprocity between a performer and simulated agents.

In terms of the social simulation, through a super agent the swarm agents perceive the performer's movement as movements of their own kind. In the world of swarms the perceivable environment is inherently contradictory because a super agent is not constrained by agents' binding social rules. What it can evoke, on the surface, is a metaphor of a conductor and an orchestra. However a conductor and orchestra share a common goal towards shaping musical events from the beginning to end given musical literature to follow. Towards progressing a coherent common goal, music performance rules are enforced a priori, cultured with shared repertoires, and manifest through unspoken collaborative agreements. Yet why should we entrain swarms to simulate human musicians when swarms already exhibit emergent patterns? Study of reciprocity aspires towards visions of future use cases, and can also function to guide us what not to do.

Prior Work in ALife-Informed Music Computation. The purpose and assessment value of the works in the field of experimental composition depart from what some people may think of as music. Therefor an experimental musical approach may depart from some research agendas in the AI community that aim to learn by modeling well-known musical styles or recognized behaviors [5,6,7,8,9,10]. Originality is also emphasized historically, for composers encouraged to learn but not to imitate previous genres. Evolutionary computing has been applied through an iterative training process to generate musical outputs including scores and recordings of synthesized sounds [11,12,13,14,15]. Bilotta and Pantano proposed a mathematical mediation to translate CA to musical language in order to transfer the semantics of complexity to a range of musical expressions [16]. The process methodology in their work comes close to structuring reciprocity between a generative mechanism and music, employing "musification

codes" as a structure that evokes aspects of reciprocity, a mathematical formalization to decode CA and re-encode them as music. The codes providing options to read a matrix frame (a grid), population growth patterns given input entropy, or local behaviors read through color or binary code. For social models, Miranda et al. applied a multi-agent listening model in composition with a genetic algorithm, given initial musical materials, feature detection and evaluation rules, and each agent adjusts its own deviation with respect to the collective outputs [17].

The above examples apply IEC – Interactive Evolutionary Computation, a process requiring extensive training and assessment to apply creative choices. Tod and Werner [18, p.319] describe this process as creating a "fitness bottleneck" and respond by applying autonomous musical agents to perform creative selections. While the process is efficient, the agents' aesthetic preferences drift without human supervision until they generate "unlistenable" music. A goal is to design a system that balances human oversight with agents' decision efficiency, which can provide exploration of emergent properties without iteration over a predetermined fitness function. Autonomous agents in flocking and swarm systems meet this requirement and methods have been implemented for efficient real-time interaction. Unemi [19] developed evolutionary efficiencies for breeding agents using a form of NIEC (Narrowly-defined IEC) [20] and Sayama [21] further optimized breeding with real-time interaction using HIEC (Hyperinteractive Evolutionary Computation).

Generative music methods extract data from simulated swarm agents to control audio signal generators. Interaction methods may be classified according to the rate and regularity of scheduled data flow between a swarm system, human users, and sound generators. Murray-Rust et al. [2006-22] and Huepe et al. [2014-23] adopt methods that apply human control during initialization only, with autonomous run-time operation dedicated to agent-agent interaction. Blackwell [2002-24], Grace [2009-25], and Davis and Karamanlis [2007-26] adopt methods to provide real-time re-initialization of swarm behavioral parameters and run-time modification of the mapping from agents' data to sound control. In these approaches the control changes are infrequent and irregular with respect to the simulation time step, and the user does not interact directly with individual agents. Other work using agents in real-time performance includes EvoMove [22-27], which applies a commensal computing scheme to provide a movement-based musical companion for dancers, with sounds conceived as ambient feedback rather than presenting a musical structure.

The above research exhibits a collective concern for management and application of emergent behaviors and methods for designing relationships between agents' tendencies and musical content. Scacher et al. [2014 CMJ-28] implement architecture for real-time performance interaction with agents, with further focus on symmetry of swarm simulation and musical structure, summarizing relationships in five classes: formal, ontogenetic, conceptual, interaction, and ecological. These are defined in terms of mapping mechanisms, classified by the extent and kind of separation between agent data and musical signal output. The concept of *designing indirection* introduced in section 2 below encompasses these functions. Social interaction between agents and human players is introduced by Choi and Bargar [2012-old38-29] and reflected by Scacher et al. as "…a shared space

within which the perceptual and behavioural properties of simulated entities and humans overlap and interrelate [op cit p. 53]." Bisig and Kocher describe performance interactions where "…human behaviour interferes with the emergent swarm behaviour and begins to form a meta-system, that might exhibit different states than the purely closed world of simulation [2012ICMC p.105-30]." The engineering of reciprocity addresses structure and design of these formations. Collectively swarm music research establishes practice of social agent simulation applied to generative music with recent examples of interactive performance. Introducing a concept of reciprocity encompasses and reframes prior research perspectives, with relevance for broader examination of reciprocity in the design and assessment of human performance with interactive social agents.

1.2   Basis for Reciprocity

The concept of reciprocity in composition relates to eliciting structure of emergent behaviors in sound rather than reproducing legacy structures or known musical styles. A mutual challenge arises to encompass the domains of performance, compositional works with evolutionary models, and the definition of composition. Ringer and Crossley-Holland [31] define composition as, "the act of conceiving a piece of music, the art of creating music, or the finished product. These meanings are interdependent and presume a tradition in which musical works exist as repeatable entities. In this sense, composition is necessarily distinct from improvisation." Here the distinction between composition and improvisation lies on the existential manifestation of the work as repeatable entity. Emergent properties contribute to interactive creation and challenge traditional boundaries that define repeatability.

The implementation of reciprocity introduces how composition is conceived for situating a performer in this work. Production of musical tones and events is inherently circular with an intimate feedback through auditory perception [32]. Kinaesthetic interaction between performer and instrument encodes a musical signal, for which the performer's movement determines the features of the signal. The choice of how to govern further movement is directed by listening to the previous states generated by previous motions. The circuit of internalization of a current state of movement, externalization of previous states to guide current state, and projection of future states to build desired hysteresis, is so highly integrated at a performer's sensorimotor level that it is difficult to separate. However the quality of performances is shaped through these fleeting moments of dual sensory-motor and cognitive processing: internalization and externalization along with consulting the musical score toward higher-level projection of future states. At the core, the chain of kinaesthetic interaction produces music, and listening shapes the chaining. These observations are experientially based supported by encouraging discoveries in neural mechanisms' temporal percepts. Recent studies suggest the brain does not process time in a serial mechanism but in distribution: interval timing is processed by feedback loops, and through these loops the output of an internal clock serves to dynamically modify its own input [33, 34]. Reciprocity is enacted by listening in this context: a performer listens and mediates the simulation with an action, utilizing

auditory feedback (sensory) to guide movements (motor), where the kinaesthetic dynamic that feeds the performer's auditory perception is extended by simulated agents' dynamics.

Reciprocity is illustrated here for two compositions with musical elements synthesized by eliciting structure of emergent swarm behaviors. Each composition constitutes two components, a score and a system of interaction. The two works, Mutandrum and Human Voice, are encoded in a computer system for programming interactive digital sound and for routing sound control data from the swarm simulation. The composer's selection and design of agents' tendencies is encoded alongside the sound synthesis design. The software encoding requires a performer's interaction to realize the compositions. For performers the compositions are represented as musical scores that indicate musical sequences but do not adopt traditional western music notation. Both scores provide blueprints for performers to work with: 1) navigation plans to move within constraints from one local free form to the next while agents play by rules at all times; 2) plans to initialize diverse types of swarms that were curated during the composition process using NIEC and HIEC. The plans are expressed in the form of scores to function as time-based musical roadmaps. The scores by no means convey all audible details or musical depth that is programmed in the composition software. The scores provide a layout of basic elements including notated instructions and a high-level rule: Move from one section to the next until all sections are exhausted. In the form of scores, at least, they structure activity so that the activity is repeatable and rehearsable. Performance here is twofold: for agents, rule-based behavioral movements; for human players, locally free form play along with agents, guided by the score. A performer does not improvise but responds to the score and to agents with situational awareness and makes decisions in a context for playing both, the composition and the swarm.

2 Rationale and Research Methods

Case studies of two compositions of this author will be presented with a comparative analysis. From experience working with nonlinear dynamical systems such as the chaotic Chua's circuit [35, 36], this author recognized that nonlinearity and self-organizational principles are often implied in a creative process itself, at subconscious level while encountering material suggestions and imageries, and at conscious level while translating them into tangible ideation. The agency of composition can be articulated by the traces of its processes and the agency of performance is articulated through realizing the composition. However other agency is challenging to articulate, such as the source of inspiration, whether it sparked as spontaneous brain chemistry or intuition. Another challenge is the management of dataflow in performance architecture to minimize a performer's cognitive overload during the complex audiovisual and sensorimotor interaction with a swarm's dynamics. To assess this complex scenario most psychoacoustics-related findings are not readily applicable, because they largely rely on isolated stimulus tests in simple settings. However, the recent discussions in neural basis of music perception [37] and the ERP processing time window [38] indicate more promising directions. orks presented here was motivated three ways: 1) to articulate

examples and implications of the hidden agency of creative choice, by introducing swarms as visibly working through an artistic inquiry; 2) to explore how swarm dynamics may yield new pathways in music creation, and further how this inquiry may contribute to the artificial life research agenda; 3) to gain quantitative insights for designing indirection for a complex evolutionary interaction, by constructing data through the case study.

2.1   Analytic Autoethnographic Method

The workflow for this research begins with a music composition process, continues to a music performance process, and culminates in an assessment process that reflects on previous stages to develop an analytical position. This workflow exhibits parallels with interaction design, then implementation, then testing. The case studies here draw upon musical work by this author as composer and performer. With awareness there may be many shortcomings in autoethnographic methodology in general [39], the methodology adapted here can be best described as an analytic autoethnography [40], including numerical data, towards establishing a new theoretical framework bridging art and science. Each step of the way keeps in sight the requisite variety [41] between the two domains with an attempt to avoid blind spots. This method recalls cybernetics of observing systems as introduced by von Foerster's analysis of self-organizing systems and their dependence upon energy and structure in their environments [42]. Self-reflexive research is performed by truly searching again one's creative results, with a reflective investigation that is objectively driven. The aim is to formalize and reframe methodologies, techniques, and aesthetic criteria, along the way noting right or wrong design intuitions, and other factors constituted for finishing the two works

Framing Research Questions. When designing a system to compose in and compose with, the creative and technical problem space hinges on non-typical conditions that juxtapose two domains: a music composition that uses a fixed notated score, and an evolutionary system that exhibits emergent behaviors. By explicating this problem the working question is framed: How do you work with a collective body that follows simple rules and self organizes, and displays unexpected patterns of behaviors, to produce and perform music that can be reproducible and rehearsed? This questions motivates a series of investigations that were undertaken roughly in the following order:

1. How swarms work:
    a. How swarms work when engaged by a player
    b. How swarms work to control sound synthesis
    c. How indirection can be designed, using knowledge gained from (a) and (b), leading to how to work with swarms

2. How to work with swarms
    a. How to work with swarms when the compositional plan is exogenous to the swarms' world – applying compositional agility for planning with adaptability

    b. How to work with swarms within an overall plan – applying compositional strategies for performance

    c. How to work with swarms in multiple levels of indirection and time scale

Through these research questions, three types of agencies are articulated: swarm, composition, and performance. The undercurrents of discussion are based on results obtained from investigations with this framework. Sections 3 and 4 apply the two framing topics, how swarms work and how to work with swarms, drawing upon comparisons of the two compositions. Section 3 introduces two techniques: emergent feature analysis and procedural sound design patterns. Section 4 surveys musical structure, discusses model-based indirection, and introduces PGAU as a tool for quantifiable analysis of performance engagement with agents. Section 5 focuses on reciprocity, and section 6 reviews the investigation of indirection and time scale, using the LIDA framework [43, 44] for comparison with temporal studies of cognitive cycles.

3   How Swarms Work in Performance

The model applied in this work is the heterogeneous swarm model developed by Sayama [21]. Unlike Reynolds' canonical flocking algorithm [45] the agents are decentralized with no leader. To initialize, typically 100 to 300 agents are set in motion with constrained random positions and initial velocities. From this initial condition, each agent's motion is dynamically influenced as it behaves with respect to other agents it detects within its perceptual range, simulating social engagement, otherwise straying in random motion. Sayama provides an efficient matrix method called a recipe, in which the number of agents and rules of behaviors can be initialized as a set of parameterized values, creating an agent behavioral type. Initializing more than one recipe for separate sets of agents forms heterogeneous swarms. Any agent that perceives another is affected in movement regardless of behavioral type.

    Data is extracted from performances with swarms using touch sensitive surfaces. Agents are visualized, a performer can touch the visual display and agents respond. Data from swarms is interpreted to generate sound control data transmitted to real-time sound synthesis engines. Data mapping includes pattern recognition and feature extraction from swarm behaviors, applying a model based design of indirection that prolongs temporal dynamics, discussed is section 4.2. Details of system configuration are reported in [46].

Visualization of Swarms. Agents are visualized by colored pixels. A mixture of colored agents represents the simulation's current state. This simplicity, that agents represent nothing more than visualized state, provides an important baseline to investigate sound with swarm dynamics in tandem with its visual presentation. The whimsical behaviors of agents spontaneously bond with others then stray, creating a constant redistribution of the visual mix, a dynamic field for emergent patterns. The evolutionary trajectory of a given type of agent can be anticipated as to its patterns over time, however the detailed trajectories and patterns of heterogeneous swarms are not predictable

beyond a local limit. Often clusters of agents forming patterns are a dominant visual feature; cluster formation is adopted for sound design to reflect this feature, focusing on tracking clusters at the cost of not tracking behavior by agent type.

Color is used differently in the two compositions to support differences in performance requirements. In Mutandrum color visualizes individual agent type and this relates to tone quality transformations (see Figure 1). In Human Voice color visualizes collective agent behavior and this relates to the number of active sound sources (see Figure 2). Visualization conveys structural differences of the two musical works, in terms of different types of sounds and different methods for agent data to transform sounds. In Mutandrum, agents' color by type highlights heterogeneous clusters and illustrates distributions of agent types. Visualizing clusters' internal structure draws attention to clusters' internal dynamics, and this data is applied to control timbre transformations (tone quality) of sounds. In Human Voice, agents' color is determined by cluster sequencing and aids the performer's control of sound by illustrating and tracking the progressive number of clusters. The number of clusters determines the number of active sound sources, a requirement specified in the performance score. To summarize: the focus on inter-agent dynamics in Mutandrum supports a timbre-dominant compositional approach, and the focus on multi-cluster formations in Human Voice supports a compositional approach based on polyphony and counterpoint.
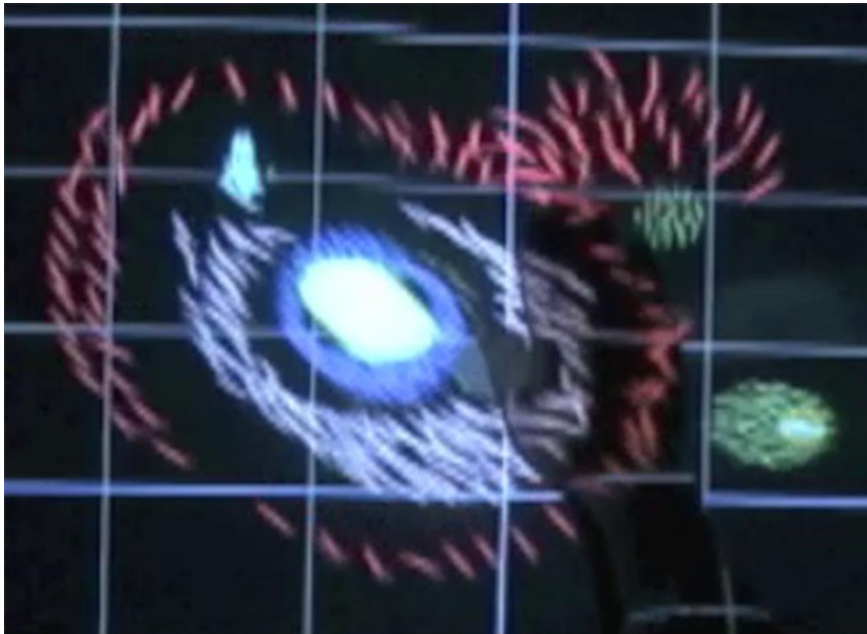

Fig. 1. Color visualization of agents by type, from *Mutandrum*. The area within each grid measures approximately three inches by four inches. Resolution constraints are discussed in section 4.1. The performers' hand is distorting the image projection to the right of center.
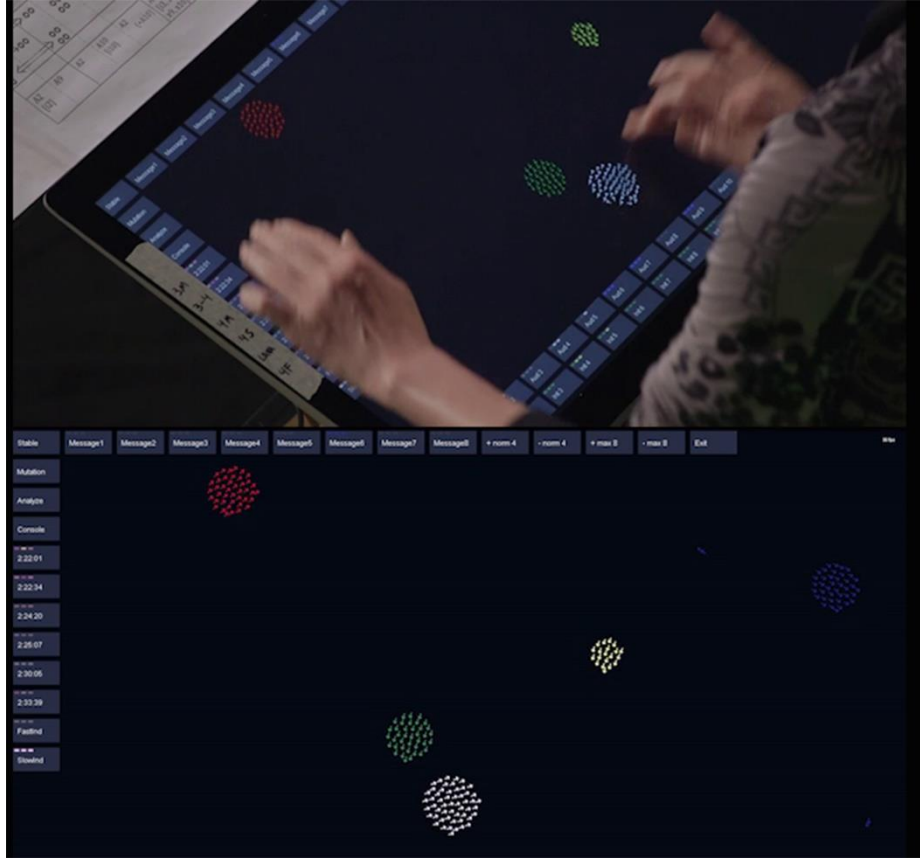
Fig. 2. Color visualization of cluster formation on a touch sensitive screen, from Human Voice, showing performance configuration.

Emergent Feature Analysis. Cluster formation is a prominent emergent feature of swarm behavior and is important for control of musical events and structures. Emergent features are not formally modeled in the swarm simulation, so to extract data of clusters and other features for sound control the performance system implements feature recognition. Table 1 presents types of swarm data grouped by duration range required for extracting the data. In Table 1 agent data AD1, AD2, and AD3 depend upon AD4 – data of cluster recognition. AD4 requires automated feature detection, as the clusters are not represented in the simulation. Emergence of these features is uncertain and influenced by a performer, therefor AD4 durations required to induce phase transitions may be prolonged beyond the values indicated in Table 1.

Table 1. Agent data grouped by duration range required for feature formation, data extraction and feature recognition of swarm state

Agent Data Type	Agent Data Source	Duration Range needed to extract feature data
AD1	Number of Agents in each Cluster	15-20ms
AD2	Energy of Agents	25-35ms
AD3	Area of each cluster	50-70ms
AD4	Position of each Cluster in performance space	75-100ms
AD5	Velocity of each Cluster	80-120ms
AD6	Cluster Deformation	250ms to 2 sec
AD7	Cluster Divide	1sec to 3sec
AD8	Cluster Merge	500ms to 2sec

Sound Design Patterns. A Sound Design Pattern (SDP) is a computational model designed to generate coherent combinations of audible attributes [46, 47]. Attributes and their transformations can be organized by duration range, combining the duration required to generate a pattern and the duration required to recognize the pattern, such as a performer responding to influence a pattern. Table 2 presents SDP types based on duration ranges of audible attributes used in SDPs. Multiple time scales of sound transformation are concurrent in a sound, and attributes of shorter duration contribute to attributes of longer duration. SDP types group sound control parameters to coordinate control data across multiple duration ranges. SDPs are not equivalent to sound sources; several SDP can combine to control a single sound source, or one SDP can contribute to the transformation of several sound sources. Combining SDPs provides a method to control features and transformations of complex sounds, and enables the use of swarm data to control transformations of attributes concurrently at multiple time scales. The lower half of Figure 6 compares the duration ranges of agents' emergent behavior and SDP audible attributes. The durations of swarm features suggests a "best fit" model by applying swarm data to SPD types of similar duration.

Table 2. Audible attributes of Sound Design Patterns grouped by Duration Range.

SDP Type	Audible Attributes of sounds	Duration Range of SDP
SDP1	Pitch change, loudness change	50ms-200ms
SDP2	Timbre, Resonance, Filtering	200ms-500ms
SDP3	Sound Source Location cues; Spatial and Directional cues	450ms-2.0 sec
SDP4	Sound Event (from onset to evolution to termination)	500ms up to many seconds
SDP5	Patterns of Rhythm, Tempo, Spoken Words, Melody	1sec up to many seonds

4 How to Work with Swarms in Experimental Performance

Traditional music composition dimensions include 1) form and tonal relationship, 2) temporal arrangement of elements including structure of repetition, 3) thematic development, 4) harmonic progression, and 5) instrumentation. Working with swarms does not deviate from these dimensions, but shifts paradigms in the ideation of the elements in them.

4.1 Experimental Compositional Dimensions

1) FORM: In both compositions, the musical form organizes types of agents' collective behavior and associates these to sounds. Behavioral types are used to organize musical structure.

2) TEMPO: In both compositions, the agents' behaviors were systemically studied for temporal signature then categorized by tempo and temporal agility: the degree of slow to fast, and the range of fluctuation in tempo caused by the degree of stability in swarm dynamics. The temporal signatures are investigated through alternating sound synthesis methods then curated as musical elements using coupled sets of parameters mapping swarm domain to sound range. To produce exact repetition a coupled set is initialized with the same initial condition.

3) THEME: Thematic development in music often relies on recognizable units of tones and rhythms. In Mutandrum repeated patterns output by tone generators were modulated by swarm dynamics, applying progressive fragmentation. In Human Voice, phrases with spoken texts functioned as semantic themes with musical meanings, which are not solely dependent on the meaning of individual texts but emerge from combinatorial bindings according to swarm dynamics.

4) PROGRESSION: Harmonic progression is a dominant technique in Western common practice tonal music. In these works the function of progression is carried by timbre (Mutandrum) and texture (Human Voice). Instead of melody and harmony, Mutandrum applies techniques for progressive derivation of spectral evolution (tone quality) in sounds, using swarms' evolutionary dynamics. Human Voice exploits additional progression in poetic form and linguistic content by generating multiple voices that present spoken texts in layers.

5) ORCHESTRATION: Instrumentation is orchestrated by sound source selection and by Sound Design Patterns (SDP), and these associations are controlled by selected behavioral tendencies of swarms; see Tables 1 and 2 and the discussion below regarding sound source selection and manipulation.

Mutandrum: Musical Goals and Features. Mutandrum explores the plasticity of sound in relationship to the plasticity of swarms, described as "sound tangibility" in [47]. This describes the kinesthetic relationship of motion dynamics between a human agent and simulated agents, and how auditory feedback confirms that relationship. The ability to engage a visualization of agents through a touch interface evokes an interactive "electro-visual-acoustic" experience, with attributes determined by relative visual

and audible responsiveness. To explore audiovisual plasticity, swarm clusters' symmetry and shape deformation are identified as top priority for feature recognition. Cluster deformation data is applied to synchronize transformations of sounds with changing cluster shapes, ensuring an obvious local correspondence between performer's action, swarm response and sound response. To further establish local correspondence the cluster's position in x-axis in the performance space drives the stereo position of the sound source associated with the cluster. Additionally the temporal dynamics of the micro-movements, or "jitters" of the agents are emphasized by sounds. When swarm jitter values are highest the sounds return strong rhythmic features and a notable pulse. When swarm jitter values are attenuated the sounds return sustained tones focusing on timbre transformations.
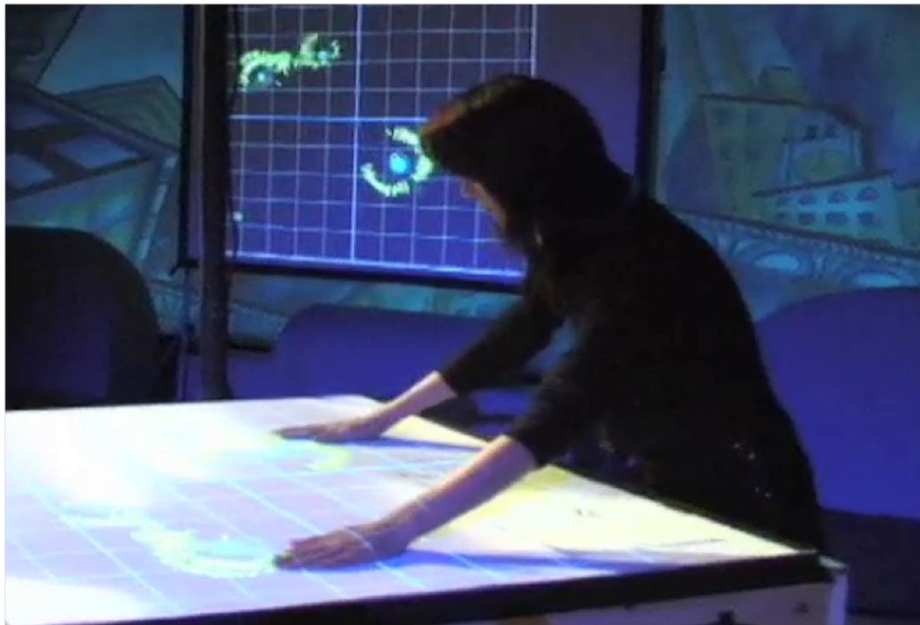

Fig. 3. Performance with large tabletop capacitive panel enabling hand-sized control regions. Visualization of swarms is projected on the performance surface from above and also projected for audience members.

Compositional choices in Mutandrum reflect the physical limitations of the capacitive surface used for the original performance, supporting only two simultaneous touch points on a 36 by 48-inch table top. The reliable resolution of position sensing for touch interaction was an area of roughly 0.3 inches square with fuzzy boundaries, approximating a differentiable resolution of 160 x 120 units. Agents visualized at 1024 x 768 pixels scaled to 48 x 36 inches producing a unit of about 0.05 inches/pixel. The ratio of this resolution to the rate of change of agents' movements and the lower-resolution capacitive sensing determines the rate of change of sound control data. Given these constraints, manipulation of large clusters is most accurate and reliable; in response

compositional design focuses on spectral explorations therefor strategizes with a limited number of sound sources. During the flow of the performance, a maximum of four complex sources are made available, and the concurrent number of sound sources varies from section to section. The quality of sounds is determined by a dominant frequency and spectral (tone quality) portrait. Cluster size determines the frequency: the larger the cluster the lower the fundamental frequency. The shape of clusters influences the spectral resonance. To avoid the feature recognition task of disambiguating complex cluster shapes, data extraction is only applied to detect relative deviation from the symmetry of the native cluster state, which is usually circular. The deformation data is applied to vary spectral resonance: the thicker the shape, the greater the resonance. All local cluster deformations dynamically affect the quality of sounds in all states, as sounds were designed using SDPs with high-level synthesis parameters exposed to receive the deformation data. Figure 3 shows a performance image with swarm visualization projected for the audience.

Human Voice: Musical Goals and Features. Voice explores a poem and uses the recorded voice of the poet as the primary sound source. The compositional intent is to create emergent discourse through structure of call and response by constructing multiple voices of the poet as sound sources. Onsets and durations of sound sources emphasize temporal dynamics at phrase level, rather than the micro-temporal levels of timbre transformation as in Mutandrum. The number of sound sources and when they come and go is paramount, with the management of multiple voices derived from cluster division, merger, and movement, evoking musical idioms such as polyphony and antiphony. In contrast to Mutandrum, the use of timbre in Human Voice is subordinated. Sounds do not undergo major timbre transformations synchronized with cluster transformations. Timbre is applied with nuance to help distinguish multiple voices and to enhance the clarity of spoken texts. Local synchronization of image and sound is not emphasized by cluster deformation, but only at the points of cluster division and merger. These cluster phase transitions are consistently reflected by the introduction and termination of voice sources. The autonomous phase transition behavior of clusters is an important contributor to the musical texture. Cluster states also require close monitoring by the performer to defer unwanted phase transitions. To do this the performer applies a "shepherding" movement technique to prolong the current swarm state and achieve longer phrase structures (see "Melisma" in Table 3).

Compositional choices in Human Voice reflect the physical affordances of a digital capacitive touch screen with a 16:9 ratio, 1920 x 1080 pixel resolution, and screen area of 21 x 12.5 inches. The resolution of 0.01 inches per pixel supports ten fingers' independent touch points capacity with arm arch range and wrist orientation supported by the performer's standing position. Figure 4 shows performance with ten-finger touch applied to swarm visualization projected for the audience. As spoken text is present throughout the piece, the compositional design focuses on textual antiphony and on polyphonic explorations. Therefor the strategy is to present these voices with intelligibility and clarity, avoiding extreme timbre and pitch deformation while applying subtle affects such as chorusing and simulated spatial characteristics such as sound source locations. To achieve this design objective the deformation relationship between cluster

and sound is not as heightened as in Mutandrum. At all times the number of sound sources are subjected to the continuous tracking of four distinct levels of cluster bifurcation. Therefor in Human Voice, swarm bifurcation has a structural role in the musical progression.
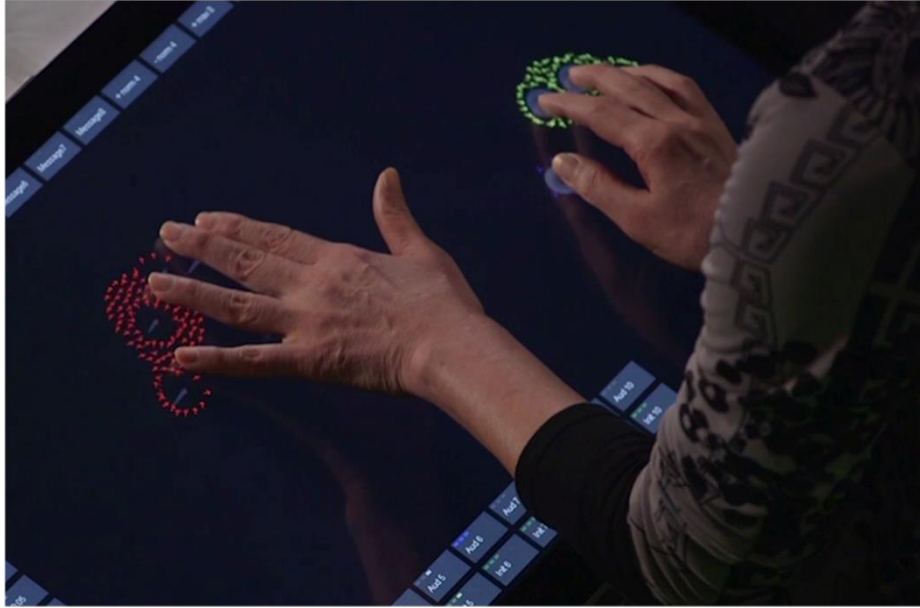

Fig. 4. Performance of Human Voice with high-resolution capacitive touch screen, showing ten-finger control signal capacity
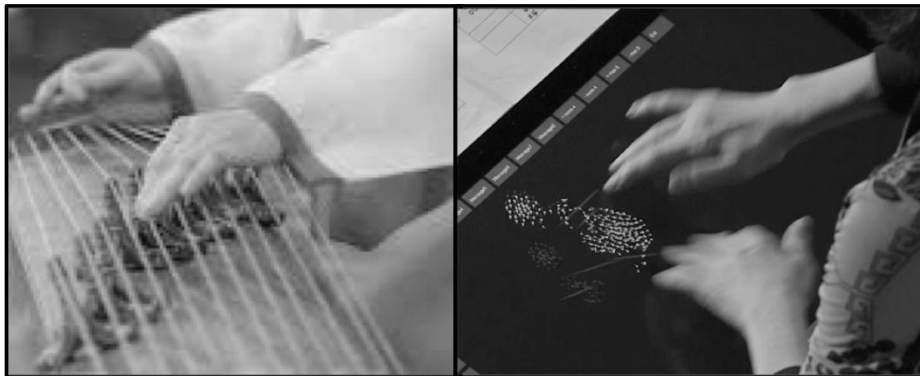

Fig. 5. Comparative illustration of a performer's kinesthetic orientation with the gayaguem (left) and swarm agents (right)

As an instrumental paradigm, there are many table top models. Touch introduces excitatory signals into the system. From the author's experience, performing Human Voice is reminiscent to the 12 silk string instrument called gayaguem, illustrated in

Figure 5. Of course, these have completely different tactile sensation and resistance; still the kinaesthetic address from a performer's perspective is highly compatible. Having performed gayageum, this author can report that cognitive plasticity of digital music instruments is possible with proper HCI engineering with adequate building blocks.

4.2 Design and Implementation of Model Based Indirection

The experimental system presents an interaction modality through which a performer inputs data to sound indirectly, based upon a model. Performer's actions on the touch screen engage swarm agents, influencing agents' collective emergent features, then control signals are transmitted from feature extracted data to parameterized sound synthesis engines. This pathway of indirection provides affordances for reciprocity between the performer and the swarm simulation. With traditional music instruments reciprocity is generated when a performer inputs excitatory energy into a resonating body, and the instrument amplifies and differentiates that energy to contribute the resulting sound quality. While learning to play music instruments entails learning to shape the excitatory input patterns, learning to play swarms in these works entails learning the model of introducing excitatory energy into sound engines by working with swarms through social engagement.

Designing indirection and engineering audio-visual presentation including process time is guided by sensitivity for intersensory asynchronies in perception, for example the Just Noticeable Different (JND) boundary of 50ms for cross modal performance and permissible ranges. The compiled study from [48] reports optimal intersensory JND between audio and visual when audio is presented 50ms before visual. Due to the brain's capacity for adaptively widening, when audio precedes visual the permissible JND ranges from 250ms, when audio follows visual it is 150ms [49, 50]. In another example, in Human Voice the most frequently used gesture articulation unit is longer than 500ms in order to accommodate syllable durations ranging 240ms to 340ms.

The analogy of HCI direct manipulation [51] in music performance comes close to instruments like harp and drum, but is not ubiquitous. Musical instruments are physically fabricated complex systems refined through extensive conditioning and tuning. The response characteristics of an instrument are complex and nonlinear: to produce musical tones requires skillful manipulations. Sound is produced by microvariations at oscillatory level; a performer cannot directly control single oscillations or manipulate all conditions of instrumental components to shape an absolute result.

Musical instruments are designed to transmit performers' nonlinear energy patterns into the physical components of the instrument that shape the quality of sound. This is why a large class of musical instruments can be modeled with nonlinear oscillators coupled to passive linear systems [52, 53]. This scenario tests the limits of the analogy of direct manipulation adopted from HCI because the music instrument is already a case for indirection for resulting sound. An obvious example is a violin. The bow is a physical model that offers an indirection with a degree of freedom to generate more variety of sounds than afforded by bare fingers. And this indirection can be computationally modeled as bow pressure and angle, etc. to simulate the friction dynamics [54, 55, 56]. For a working definition, model based indirection is a technique to instrument

an extended interactive pathway to optimize the relationship between performance input and resulting sounds, where the extension model engages a generative mechanism.

4.3   Performance Gestural Articulation Units (PGAU)

Table 3 presents a comparative analysis of indirection and resulting temporal dynamics in the two compositions. Indirection is observed with respect to the actions of a performer and the corresponding swarm data and resulting sound patterns. The table compares indirection across four types of hand movements, referred to as Performance Gestural Articulation Units (PGAU). In order of increasing duration and indirection the units are named after Gregorian notation, punctum, brevis, longa, and melisma. The nomenclature is adopted in reference to chant tradition where durations are context dependent with no absolute values. For a generalizable definition applicable to both human movements and computational model, 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. PGAU are applied to express a functional contribution to the temporal dynamics in situ of performance along with system constituents [47]. In this context, the constituents are swarms, SDPs and a performance score.

For each type of PGAU in the two compositions, Table 3 shows SDP models and agent data (from Tables 1 and 2). The two compositions adopt different combinations of agent data groups and SDPs for PGAU. The differences reflect the different sound sources and compositional structure in the two cases. As an example, Human Voice does not apply SDP1 and SDP2, which generate micro-variations of pitch, loudness, and timbre. This is because in Human Voice, these attributes are rendered by the recorded voice performance of a poet, and are preserved at the performance timescale rather than transformed. Timbre and loudness variation are applied using longer–duration SDPs.

PGAU is an expression unit of temporally defined hand gestures analogous to excitatory input of a music instrument. Similarly, impact from PGAU propagates through indirection dataflow in the system, accumulating data extraction time and sound processing time while inheriting the temporal definitions of the SDPs. These accumulations result in duration of indirection extending sound events beyond PGAU hand movement. Table 3 illustrates how PGAU engages simultaneously multiple agent data types and SDPs, generating multiple levels of indirection. Indirection built into the design is most easily observed in the latency of prolonged sound responses to PGAUs. The Maximum Latency of Indirection (Table 3) was comparatively observed in recorded performances of each composition, measured 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. While a novice can play with this system, a performance skill deepens in part through mastering the relationship between PGAU, swarms, and sounds shadowing swarms through model based indirection.

Table 3. Comparing Indirection Profiles of two compositions, *Mutandrum* and *Human Voice*. Comparison is drawn against the duration ranges of performance gestural articulation units (PGAU) and latency of indirection for each PGAU. Duration ranges of agent data types are presented in Table 1. Duration ranges of sound design patterns (SDP) are presented in Table 2.

PGAU:	*Mutandrum*				*Human Voice*			
	PGAU Duration Range		Agent Data Type		SDP Type		Maximum Latency of Indirection	
Punctum	100ms-500ms	300ms -1sec	AD 1, 2, 3	AD 6, 7, 8	SDP 1 & 2	SDP 4	500ms	1.5 sec
Brevis	300ms - 2sec	500ms - 2sec	AD 3 to 6	AD 7 & 8	SDP 1,2, 3	SDP 3 & 4	1sec	3 sec
Longa	4sec - 9sec	3sec - 15sec	AD 4 to 8	AD 4, 5, 6	SDP 2, 3, 4	SDP 3, 4, 5	3sec	5 sec
Melisma	8sec - 13sec	5sec - 60sec	AD 4, 5, 6	AD 1 to 5	SDP 2, 3, 5	SDP 3, 4, 5	4sec	6 sec

5 Composing Structured Reciprocity

Each composition encodes association of temporal dynamics between swarms and sounds, as follows. For each section of the musical score an HIEC selection of swarm agent types has been assigned, and for each section of the music performance code SDPs have been developed to respond to agent behaviors. The reproducibility of these associations enables a performer's learning curve when encountering the predicted behaviors so that each composition can be rehearsable. Reciprocity is structured in a time window that represents concurrent presence of agent behaviors and SDPs. This structure is composed anticipating the efforts require to rehearse, to perform, and to listen. Emergent swarm behaviors generate control data for sounds' attributes that are too detailed for a performer to control directly. The performer rehearses the reliability of swarm behaviors and how to induce and prolong cluster phase transitions and steady states, generating sound control data required to render the musical score. Reciprocity emerges as a practice of shared control data generation in performance.

The following outlines the process end-to-end for encoding, activating, generating, and measuring reciprocity. Encoding: In the experimental system reciprocity is encoded in the temporal dynamics of the indirection model. This encoding combines the model of PGAU performance engagement, the swarm feature data extraction, and the temporal dynamics of the SDPs. Activating: Reciprocity is activated by the performer's PGAU prioritized for swarms' behavioral tendencies and the levels of sound transformation specified in the score. Generating: Reciprocity is generated by the requirements of the musical score and the performer's engagement with the swarm tendency "envelope"— the average rate a swarm recovers its native tendency following a PGAU intervention. Measuring: Reciprocity is measured by the performer in judging the need to induce

changes in swarm behaviors with respect to the SDP. During the bidirectional process when the composer develops SDPs and swarm tendencies, the composer anticipates the performer's measure of reciprocity, as follows: 1) SDPs are computationally modeled to express the compositional idea; 2) swarms are designed to exhibit tendencies that are a good fit for expressive control of the SDPs; 3) swarm data feature extraction is tuned to recognize collective behaviors; 4) SDP mapping is fine tuned to reflect the variety of detail in the swarm data.
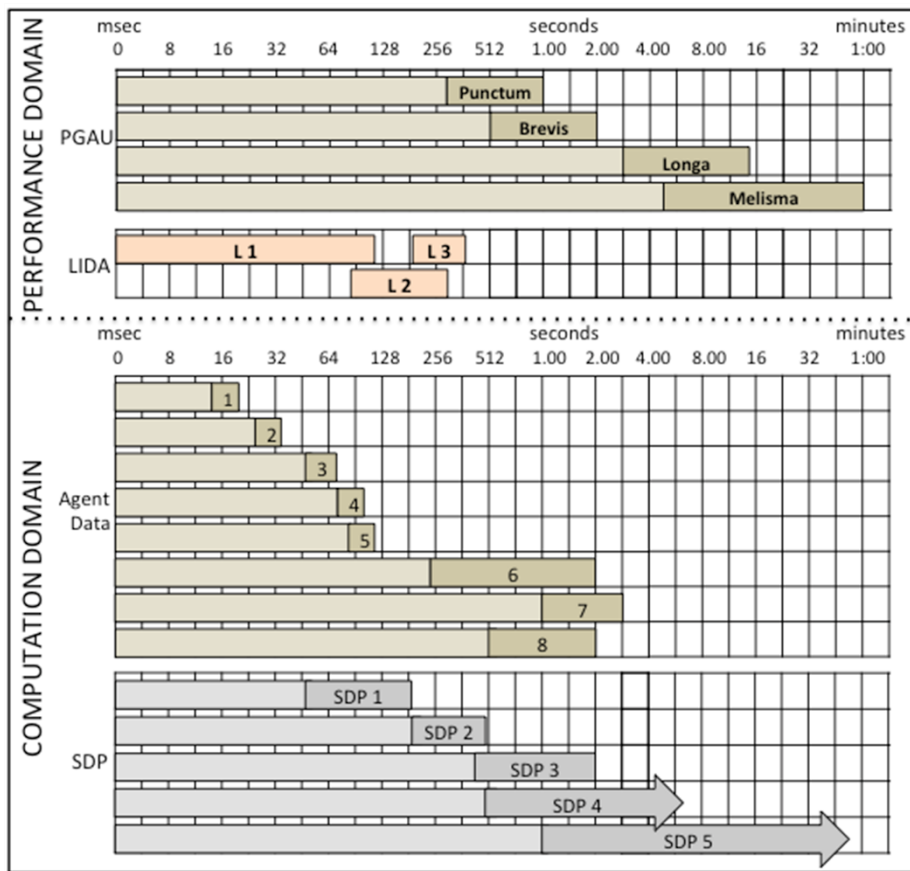

Fig. 6. Comparison of minimum and maximum duration ranges of agent data types, SDPs, PGAU types, and the LIDA model. The darker shaded areas are the valid range for each item. The duration scale left to right is Log2. Arrows on SDP4 and SP5 indicate potential prolongation.

Swarms and sounds are scheduled by a series of initializations, some automated, some activated by the performer, as outlined in a performance score. Layers of SDPs respond concurrently to parallel data streams. Figure 6 synthesizes PGAUs with the durations of four sets of agent data types (AD1-4) and five types of SDPs (SDP1-5) from Tables 1-3. The onset of a PGAU can transform multiple concurrent data streams. In Figure 3, agent data streams AD1 and AD2 are updated roughly at 10 Hz while behaviors in AD3

and AD4 emerge over longer durations. The arrows at AD3 and AD4 in Figure 6 indicate the potential for a performer to prolong those behaviors. Concurrently, SDP1 updates at about 5 Hz and SDP 2 updates at 2 to 4 Hz, while SDP3, 4 and 5 progress across longer durations. The arrows at SDP4 and 5 indicate the potential for a performer to further prolong PGAUs. Figure 6 across the top indicates maximum and minimum durations for PGAU types observed in performances of two compositions, to the duration limit of the figure.

For experimental analysis the LIDA model [36] provides a reference to measure the range of performer's engagement in reciprocity with respect to the system temporal dynamics. Figure 6 compares the duration of the LIDA cognitive cycle to durations of agent data types and SDPs, discussed in the next section. The layered profile in Figure 6 illustrates a multi-temporal capacity for designing indirection, which is used to prioritize the data from selected swarm tendencies. The mapping from swarm to sound at multiple time scales formalizes the mappings from swarm emergent features to SDPs.

6   Discussion

Reciprocity structured by designed indirection in the signal path enables the integrity of the three agents of co-creation: swarms, composition, and performance. Designing indirection utilizes PGAU, swarms' behavioral tendencies, and SDP. Coherently designed, structured reciprocity leverages performer's associative memory for learning and refining performance skills with the interactive system.

During the performance with audiovisual continuity driven by swarm dynamics, a performer experiences a time window for action selection across evolving swarms and a stream of sounds. SDPs are building blocks for model based indirection that increase the degree of freedom and capacity for PGAU, to act in nested time windows with respect to the concurrent duration groups (Figure 6 and Tables 1 and 3), with multiple layers of indirection extending from the swarm state. The swarm data at any given time will be reflected in sound during the time delay of each SDP window. The temporal extension of each indirection layer depends on the SDP parameters and swarm tendency envelope. For understanding performers' parallel engagement with multiple temporal layers, LIDA's single cognitive cycle provides a reference model to distinguish perception, understanding, and action selection.

The LIDA model provides a 260-390ms cognitive cycle with an initial 200-280ms unconscious processing comprised of perception (80-100ms), then understanding, followed by conscious action selection (60-110ms). The durations considered in composing for reciprocity reflect this model, commencing with 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 within a perception time window of 50-100ms (compared to 80-100ms in LIDA). SDP2 events occur from 200-500ms, many within the 200-280ms LIDA unconscious period. Roughly a 50-280ms time windows encompass the basic perception of pitch, loudness, and pulses, and continuing transformation of sounds that have been already set in motion.

A performer's action decision falls into two categories at any given moment, PGAU onset and the purpose. Most of the time, performers intervene in bifurcation modes or influence the transformation modes of already launched events. SDP1 and SDP2 do not require interpretation and planned response and can be modified by performer's ongoing actions. Recognition of spatial cues in SDP3 falls on the borderline of the 390ms cognitive cycle proposed in the LIDA model.

Agent data is detected in multiple time scales: 15-120ms for AD1 and AD2 afford duration for sensorimotor response; 250ms-3 sec for AD3 and AD4 afford duration for recognition of evolving patterns and phase transitions. SDP4 can carry all of SDP1, 2, 3 in a single sound event and SDP5 can carry multiple sound events forming a larger phrase structure. SDP4 and SDP5 are programmed to host indirection with durations greater than the 390ms action cycle. The duration required for action planning with SDP4 or SDP5 depends on the emerging complexity and temporal dynamics, but is approximately twice the minimum duration (500ms) of SDP4, which allows at least two cognitive cycles for action planning.

To conclude: 1) Structured reciprocity helps optimize the use of evolutionary tendencies including their intuitive qualities, towards desired outcomes in interactive applications. The ongoing generative mechanism of the swarm model accompanies performers with a "visual metronome" to gauge and anticipate tendencies and emerging patterns. The intuitive quality of visualized swarm dynamics aids the performer's capacity to predict future states and to plan action repertoire towards future states of sounds. 2) Observational data from the case study lends quantitative guidance towards formalizing indirection for complex interactive scenarios. Model-based indirection provides performers with time latency needed for action planning. With indirection design, the goal is to engage swarms in the model of indirection, which facilitates the efficient use of data from swarms' inherent tendencies. The efficiency is achieved by consulting the swarms' generative mechanism to drive details of musical outcomes along with SDPs. The better the fit of inherent swarm tendencies, the closer the resulting association of visualized swarm dynamics with anticipated sounds.

References

1. Galanter, P.: Computational aesthetic evaluation: past and future. *Computers and Creativity*. Springer Berlin: Heidelberg, 255-293 2012.
2. Boden, M, and Edmonds, E.: What is Generative Art? *Digital Creativity* **20** (1/2): 21–46 (2009).
3. Stiny, G. and Gips, J.: Shape Grammars and the Generative Specification of Painting and Sculpture. In: Freiman, C. V. (ed.) *Information Processing 71,* Amsterdam: North Holland, 1460-1465 (1972).
4. Romero, J. J. and Machado, P: T*he Art of Artificial Evolution.* Berlin: Springer (2008).
5. Cope, D.: *Computers and Musical Style*. A-R Editions, Madison (1991).

6. Dubnov, S., Argamon, S. and Burns, K.: *The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning.* Springer SBM (2010).

7. Todd, P. M.: A Connectionist Approach to Algorithmic Composition. *Computer Music Journal* 13(4), 27-43 (1989).

8. Brooks, F., Hopkins, A., Neumann, P. & Wright, W.: An experiment in musical composition. *IRE Transactions on Electronic Computers*, Vol. 6, No. 1, 175-182 (1957).

9. Self Organizing Maps clusters arbitrary data presented to the input layer without feedback. Phon-Amnuaisuk, S.: Evolving music generation with SOM-fitness genetic programming. In: *Lecture Notes in Computer Science*, 4448 LNCS, 557-566 (2007).

10. Law, E. and Phon-Amnuaisuk, S.: Towards Music Fitness Evaluation with the Hierarchical SOM Applications of Evolutionary Computing. In: *Lecture Notes in Computer Science*, 4974 LNCS, 443-452 (2008).

11. McDermott, J., Griffith, N. and O'Neill, M.: Toward User-Directed Evolution of Sound Synthesis Parameters. In: *Applications of Evolutionary Computing*, Proceedings 3449, 517- 526 (2005).

12. Mitchell, T. and Pipe, A.: Convergence Synthesis of Dynamic Frequency Modulation Tones Using an Evolution Strategy. In: *Applications on Evolutionary Computing.* Lausanne: Springer, 533-538 (2005).

13. Magnus, C.: Evolutionary Musique Concrète. In: F. Rothlauf & J. Branke (eds.) *Applications of Evolutionary Computing*, EvoWorkshops Berlin:Springer, 688-695 (2006).

14. Hazan, A., Ramirez, R., Maestre, E., Perez, A., and Pertusa, A.: Modelling Expressive Performance: A Regression Tree Approach Based on Strongly Typed Genetic Programming. In: *Proceedings of the European Workshop on Evolutionary Music and Art*, Budapest, 676-687 (2006).

15. Ramirez, R., Hazan, A., Marine, J. and Serra, X. Evolutionary Computing for Expressive Music Performance. In: Romero, J. and Machado, P. (eds.), *The Art of Artificial Evolution.* Berlin Heidelberg: Springer-Verlag, 123-144 (2008).

16. Bilotta, E. and Pantano, P.: "Artificial Life Music Tells of Complexity." In *Proceedings of Artificial Life Models for Musical Applications ECAL 2001 Workshop*, Prague (2001).

17. Miranda, E., Kirke, A. and Zhang, Q.; Artificial Evolution of Expressive Performance of Music: An Imitative Multi-Agent Systems Approach. *Computer Music Journal* 34(1): 80-96 (2010).

18. Todd, P. M., & Werner, G. M.: Frankensteinian Methods for Evolutionary Music Composition. In: Griffith, N. and Todd P. (eds.) *Musical networks: Parallel distributed perception and performance*. Cambridge: MIT Press/Bradford Books (1998).

19. Unemi, T.: Simulated breeding–a framework of breeding artifacts on the computer. *Kybernetes*, 32(1/2):203–220 (2003).

20. Takagi, H.: Interactive evolutionary computation: Fusion of the capabilities of optimization and human evaluation. *Proceedings of the IEEE*, 89(9):1275–1296 (2001).

21. Sayama, H.: Decentralized Control and Interactive Design Methods for Large-Scale Heterogeneous Self-organizing Swarms. In: Almeida e Costa, F. (ed.), *Advances in Artificial Life. ECAL 2007,* LNAI 4648, Berlin: Springer-Verlag, 675–684 (2007).

22. David Murray-Rust, D., Smaill, A. and Edwards, M. MAMA: An architecture for interactive musical agents. In: G. Brewka et al. (Eds.) *Proceedings of ECAI 2006*. IOS Press (2006).

23. Huepe C., Colasso M., and Cádiz R.F.: Generating Music from Flocking Dynamics. In: LaViers A. and Egerstedt M. (eds), *Controls and Art*. Springer. DOI: 1010.1007/978-3-319-03904-6_7 (2014).

24. Blackwell, T.: Improvised music with swarms. In: *Proceedings of the 2002 Congress on Evolutionary Computation, CEC '02*. DOI: 10.1109/CEC.2002.1004458 (2002).

25. Grace, L. D.: Music box: composing and performing visual music. In: *Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE '09)*. ACM, New York, NY, USA, 445-445. DOI: 10.1145/1690388.1690493 (2009).

26. Davis T. and Karamanlis, O.: Gestural Control of Sonic Swarms: Composing with Grouped Sound Objects. In: *Proceedings SMC'07, 4th Sound and Music Computing Conference* (2007).

27. Peignier, S., Abernot, J. Rigotti, C. and Beslon, G.: EvoMove: Evolutionary-based living musical companion. *Proceedings of 14th European Conference on Artificial Life* (2017).

28. Schacher, J., Bisig, D. and Kocher, P.: The Map and the Flock: Emergence in Mapping with Swarm Algorithms. *Computer Music Journal*, Vol. 38, No. 3, 49-63. DOI: 1010.1162/COMJ_a_00256 (2014).

29. Choi, I. and Bargar, R.: A Playable Evolutionary Interface for Performance and Social Engagement. In: Camurri, A. and Costa, C. (eds.) *Proceedings 4th International ICST Conference on Intelligent Technologies for Interactive Entertainment*, LNICST 78, ISSN 1867-8211, Heidelberg: Springer, 170-182 (2012).

30. Bisig, D. and Kocher, P.: Tools and Abstractions for Swarm Based Music and Art. In: *Proceedings of 38th Annual International Computer Music Conference, ICMC 2012,* Ljubljana: ICMA (2012).

31. Ringer, A. and Crossley-Holland, P.: Musical Composition. In: *Encyclopaedia Britannica* (2008).

32. Choi, I.: A Manifold Interface for Kinesthetic Notation in High-Dimensional Systems. In: Battier, M. and Wanderly, M. (eds.) *Trends in Gestural Control of Music.* Paris, IRCAM (2000).

33. Matell, M., Meck, W. and Nicolelis, M.: Integration of Behavior and Timing: Anatomically Separate Systems or Distributed Processing? In Meck, W. (ed.), *Functional and Neural Mechanisms of Interval Timing.* Boca Raton: CRC Press, 371-391 (2003).

34. Buhusi, C.: Associative and Temporal Learning: New Directions. *Behavioural Processes* 101: 1-3 (2014).

35. Choi, I.: Interactive exploration of a chaotic oscillator for generating musical signals in real-time concert performance. *Journal of the Franklin Institute*, vol. 331B nr. 6, 785-818 (1994).

36. Chua, L.: Local Activity is the Origin of Complexity. *Intn'l. Journal of Bifurcation and Chaos* 15, nr. 11: 3435-3456 (2005).

37. Koelsch, S.: Toward a Neural Basis of Music Perception – A Review and Updated Model. *Frontiers in Psychology*, *2*, 110 (2011).

38. Zioga, I., Di Bernardi Luft, C., and Bhattacharya, J.: Musical training shapes neural responses to melodic and prosodic expectation. *Brain Research*, 1650, 267–282 (2016).

39. Ellis, C., Adams, T. and Bochner, A.: Autoethnography: An Overview. *Forum: Qualitative Social Research*, 12(1), Art. 10 (2010).

40. Ellingson, L and Ellis, C.: Autoethnography as constructionist project. In: Holstein, J. A. and Gubrium, J. F. (eds.) *Handbook of Constructionist Research*, New York: Guilford Press, 445-466 (2008).

41. Ashby, W. R.: Variety, Constraint, and the Law of Requisite Variety. In: *Modern Systems Research for the Behavioral Scientist.* Aldine, Chicago (1968).

42. von Foerster, H.: Chapter 1. On Self-Organizing Systems and Their Environments. In: *Observing Systems,* Seaside, CA: Intersystems (1981).

43. Franklin, S. and Patterson, F.: The LIDA Architecture: Adding New Modes of Learning to an Intelligent, Autonomous, Software Agent. *IDPT-2006 Proceedings (Integrated Design and Process Technology),* Society for Design and Process Science (2006).

44. Madl, T., Baars, B.J. and Franklin, S.: The Timing of the Cognitive Cycle. *PLoS ONE* 6(4): e14803. doi:10.1371/journal.pone.0014803 (2011).

45. Reynolds, C.: Flocks, herds and schools: A distributed behavioral model. In *SIGGRAPH '87: Proceedings of the 14th annual conference on computer graphics and interactive techniques,*. Association for Computing Machinery, 25-34 (1987).

46. Choi, I. and Bargar, R.: Sounds Shadowing Agents Generating Audible Features from Emergent Behaviors. In: Sayama, H. et al. (eds.) *Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, Cambridge: MIT Press, 726-733 (2014).

47. Choi, I.: Interactive Composition and Performance Framework with Evolutionary Computing. In: *Proceedings of 43rd Annual International Computer Music Conference,* Shanghai: ICMA, 351-356 (2017).

48. Keetels M. and Vroomen J.: Perception of Synchrony between the Senses. In: Murray, M. and Wallace, M. (eds.) *The Neural Bases of Multisensory Processes*, Boca Raton: CRC Press, Chapter 9 (2012).

49. Besle, J., Fort, A., Delpuech, C., and Giard, M.-H.: Bimodal speech: early suppressive visual effects in human auditory cortex. *The European Journal of Neuroscience 20*(8): 2225–2234 (2004).

50. Flinker, A., Korzeniewska, A., Shestyuk, A. Y., Franaszczuk, P. J., Dronkers, N. F., Knight, R. T., and Crone, N. E.: Redefining the role of Broca's area in speech. *Proceedings of National Academy of Sciences of the United States of America*, *112*(9): 2871–2875 (2015).

51. Shneiderman, B.: Direct manipulation: a step beyond programming languages. *IEEE Computer* 16(8): 57-69 (1983).

52. McIntyre, M.E., Schumacher, R.T. and Woodhouse, J.: On the Oscillations of Musical Instruments. *Journal of the Acoustical Society of America* 74 (5), (1983).

53. Rodet, X.: Nonlinear Oscillator Models of Musical Instrument Excitation. In: *Proceedings of the International Computer Music Conference*, San Francisco: ICMA, 412-413 (1992).

54. McIntyre, M. E. and Woodhouse, J.: On the fundamentals of bowed-string dynamics. *Acustica* 43: 93–108 (1979).

55. Woodhouse, J.: On the stability of bowed string motion. *Acustica* 80: 58–72 (1994).

56. Debut, V., Antunes, J. and Inácio, O.: Linear modal stability analysis of bowed-strings. *Journal of the Acoustical Society of America* 141(3): 2107-2120 (2017).