**Intergestura: a gestural agent based on sonic meditation**
**Kieran Maraj**

## Introduction

Digital technology has complicated the notion of what a new instrument can be. Modern digital technologies have changed the way we operate in all areas of musical practice, from making instruments, to composing, performing, improvising, and listening (Magnusson 33). The lines of what can or should be considered an instrument or a system have been blurred, the boundaries between them have become porous. Contemporary research has seen creative agency granted to computers and these distinctions further complicated.

This paper presents the *Interdependence Gestural Agent (Intergestura)* system. It is an agent based system based on the works of composer Pauline Oliveros, intended as an interactive compositional and improvisational system. *Intergestura* explores musical metacreation and acts of human-machine co-creation and partnership.

## Related Work

### *Sonic Meditations*

Drawing influence from various religious meditation traditions, both Eastern and Western, American composer Pauline Oliveros published her first set of secular *Sonic Meditations* in 1974 (Oliveros 1974). The meditations are a collection of text-based scores meant for performance by musicians and non-musicians alike. Through sound and hearing, they are ways of training, coordinating and synchronizing attention and awareness (Oliveros 1984, 138). She describes attention as “narrow, pointed and selective” whereas awareness is “broad, diffuse and inclusive” (139). The enhancement and development of aural attention and awareness are some of the explicit goals of the *Meditations* project - she describes an effect of the *Meditations* as being a narrowing and focusing of attention, whether to the external world or internal bodily world, and an expansion of awareness, keeping the two in balance with each other. Observation of the two is a key part of the *Meditations* and she states: “the observation of attention and awareness is also *having* attention and awareness” (152).

The development of *Sonic Meditations* was also a highly collaborative and interdisciplinary process. In a long-term collaboration with an ensemble of musician and non-musician participants, Oliveros describes bringing in a kinesiologist/choreographer, psychologist, T’ai Chi master, and medical researchers to help with experimentation and early development of the *Meditations* (158). For example, her work with T’ai Chi master Al Chung Liang Huang brought a strong focus on the breath cycle to the *Mediations,* as well as to Oliveros’ own performance practice. As their development blurred disciplinary lines for those involved, performances of the various *Meditations* similarly blur the lines between performer and audience. In performance, all persons present are meant to take part in the piece - audience is performer and performer is audience. The distinction between specialist and non-specialist is dissolved through this process of participatory co-creation. In this way, “Oliveros is more interested in the social, psychological and even physiological aspects of music making than in its product” (Gioti 27).

There are a number of sonically interactive pieces that either focus on or seek to augment meditation practices more generally. One of the earliest examples of such is Alvin Lucier’s 1965 piece, *Music for Solo Performer*. In this piece, the performer, with help of an assistant, has a set of electrodes placed on their head for electroencephalography (EEG). The EEG signal is sent through various analog amplifiers and filters in an attempt to isolate the alpha brain waves of the performer (Straebel and Thoben 22). The resultant signals, with frequencies below the range of human hearing, are mixed and routed by another assistant and sent out to loudspeakers which are carefully placed next to various percussion instruments. The amplified alpha waves of the performer cause the instruments to resonate. Alpha waves are believed to occur when there is a lack of information processing in certain areas of the brain (Straebel and Thoben 22). Consequently, performances of *Music for Solo Performer*, which typically lasts around 40 minutes, featured the performer at centre stage, eyes closed, trying to clear their mind - a practice similar to certain forms of meditation. John Tilbury, an early performer of the piece remarks: "I think I was alone on stage for around 40 minutes trying to achieve *Samadhi* (emancipation from thought)" (Straebel and Thoben 19).

An interactive augmentation of meditation practice can be found in the works of Jiayue Cecilia Wu. Wu terms her practice "embodied sonic meditation" (Wu 307) and draws on the traditional spiritual and meditational practices of Tibetan Buddhism, as well as the works of Oliveros and composers Éliane Radigue and Phillip Glass. In embodied sonic meditations, Wu designs digital music instruments that are intended to "[work] seamlessly with the performer and [augment] the performer's voice or generated-sound so well that it becomes the performer's extended body/voice" (Wu 311). One such embodied sonic meditation is the *Tibetan Singing Prayer Wheel (TSPW)*. The piece features a prayer wheel, a traditional Tibetan spiritual tool, that has been augmented with various motion sensors, as well as vocal input. The movement of the augmented prayer wheel controls a virtual Tibetan singing bowl and additionally processes the real-time vocal input. The *TSPW* has been used to augment performances of traditional Tibetan Buddhist meditations, as well as in participatory installation contexts where "[use of] real-time body movement to manipulate sound helps laypersons, especially children, to experience sound in an intimate way, thus enhancing their embodied sonic meditation experience" (Wu 315).

Finally, a piece that explicitly augments and expands upon one of Oliveros' *Sonic Meditations* is *dispersion.eLabOrate* by Rory Hoy and Doug Van Nort. This piece features a room-scale ecosystemic augmentation of Oliveros' *Tuning Meditation* (Hoy and Van Nort 320). The *Tuning Meditation* is structured as a sort of call and response piece, with participants instructed to alternate between listening to the collective group and sounding with the group according to various rules. In *dispersion.eLabOrate,* a group of human participants is joined by a room-scale agent in performing the meditation. Just as human performers listen to and sound with the collective, the system does the same through an array of microphones and loudspeakers. It listens to the collective through the microphone array, analyzes what it hears in software, and then generates and plays into the room sounds according to the instructions of the *Tuning Meditation* - in this way, the system bears witness to human activity and augments it. Just as the original *Sonic Meditations* are co-created through a process of boundary blurring, *dispersion.eLabOrate* is co-created by and blurs boundaries between human participants and the technological system.

*Musical Agents and Co-Creation*

One of the earliest musical automata appeared in the seventh century - al Jazari's water clock, a sort of hydraulic-powered cuckoo clock featuring drums, cymbals, and trumpets (Tatar and Pasquier 56). Various types of musical automata continued to be developed through the industrial revolution - one such invention being Johann Nepomuk Mälzel's panharmonicon. Using pinned cylinders in place of punched paper rolls, Mälzel's panharmonicon functioned similarly to player pianos. The primary difference being that the panharmonicon was an orchestral machine made of constituent instrumental pieces; an early version contained multiple woodwinds and brass horns, each tuned to a fixed note and attached to mechanical actuators to produce sound. Later versions would go on to contain access to all the timbres of the classical orchestra, additionally including string and percussion instruments (De Souza 163). These early musical machines can be seen as part of a lineage that includes modern musical software agents.

Software agents are computer programs that work on user-specified tasks and can have different levels of persistence, independence, communication, and collaboration with other software agents or humans (Whalley 156). Musical agents then are simply software agents that carry out musical tasks and are a subset of the Multi-Agent System (MAS) field of study. There are a wide range of tasks and a number of ways in which they can be achieved. Tasks might include real-time melody or rhythm generation, narrowing the available note choices for a performer, or freely improvising with other software agents or human performers. Offline tasks might include things like making suggestions for instrumental parts or direction in the composition process. Agents might accomplish these tasks by interacting with rules or process-based approaches. They might have access to real-world data, like the sound produced by an ensemble of human performers, or they might exist purely within a simulated virtual environment that governs how the agents behave with one another. A MAS might be able to operate completely independently or might rely upon outside human input and interaction.

One early example of an agent-based performance system is George Lewis' 1988 *Voyager* system. *Voyager* is a software system with roots in Black American and African diasporic cultures intended for use in real-time improvisation settings that Lewis describes as "a 'program,' a 'system,' and a 'composition'"(Lewis 33). The system can perform completely on its own or it can perform with other human players, "listening" to them either through a MIDI interface or a machine listening algorithm. As part of his goal in "de-instrumentalizing the computer" (36), Lewis intends for all interaction with the system to happen through sound - there are no knobs, sliders, foot pedals or other ways for a human to interact with the system other than through sound. Like the panharmonicon, *Voyager* is made up of many constituent elements. Lewis conceives of these elements as “players” in an improvising orchestra, but they could equally be thought of as individual agents in a MAS. The multiple players are combined into ensembles with each ensemble being assigned a specific task. Players are recombined into new ensembles and tasks are changed in 5 to 7 second intervals. Tasks include melody generation, choice of pitches, assigning volume range, rhythm, tempo, probability of a note happening, and interval range, amongst others. The system can also choose whether it wants to listen to any other players or not. Lewis states, “the system is designed to avoid the kind of uniformity where the same kind of input routinely leads to the same result” (36). The system is built to be full of surprise.

Lewis seems to view *Voyager* as its own entity. He tries to understand interactivity with *Voyager* as a process of dialogue, rather than one of information retrieval. He repeats that the system is non-hierarchical - in performance, it can follow, or act as leader, it can choose to respond or not, just as a human improviser might be faced with making the same set of choices. It seems that to play with *Voyager* is to engage in an act of human-machine co-creation.

Andrew Brown’s *Ripples* is a more recent system for human-machine co-creation. *Ripples* is described as “an interactive music composition that includes a generative software system designed to be a musical duet with a live musician” and is a creativity support tool and meta-creative system (Brown 2016, 1). The system contains three instances of the same agent. These agents consist of a generative process that involves the use of a varying low frequency oscillator to generate melodic arpeggiation patterns. The agents can run on their own, but “a key element to *Ripples* is the way in which it moves beyond generation to embrace interaction with the intention of forming a cybernetic system with the musician" (Brown 2). The system "listens" to input from a MIDI interface and maintains a running memory of the most recently heard pitches and their velocities and uses this data to condition the output of the three agents. Brown writes that the interaction with the system "can be considered a micro-level call and response, but the overlapping nature and metric independence of the agents means that the effect is more of a polyphonic lattice than a distinct call and response" (2). While Brown seems to view the system less as a creative partner and instead more as a means of extending the possible creative capabilities of human performer, the human-machine interaction still remains one of co-creation - without both constituents, MAS and human musician, the creative possibilities would be reduced.

*Musical Metacreation and Agent Typology*

Musical Metacreation (MuMe) is the field of study of "partial or complete automation of musical tasks… as an interdisciplinary field, [it] is inclusive of all approaches, studies, domains and practices that automatise musical tasks" (Tatar and Pasquier 62). MuMe is the combination of the artistic field of generative music and the scientific field of computational creativity. Musical agents and MAS systems are very much a part of the MuMe field. Tatar and Pasquier write, "software agents in MAS are autonomous pieces of software which contain perception and action abilities. Applications of MAS are beneficial to modelling and designing musical creativity because musical creativity involves distributed, coordinated entities with perception and action abilities" (56).

To this end, Tatar and Pasquier propose a 9-dimensional typology of musical agents (Figure 1). These dimensions are: agent architecture, musical tasks, number of agents, number of agent roles, environment, corpus, input/output, communication, and human interaction modality.

The agent systems surveyed to create this typology largely function through the use of audio, symbolic music representation, or a combination of the two. Given that *Intergestura* uses gestures as its primary unit of interaction, which contain either solely control data or a combination of control data and sonic data, it somewhat complicates this typology and doesn't always fit neatly into the defined subcategories. However, the 9 dimensions of the typology are still useful and worth keeping in mind in describing the system going forward.

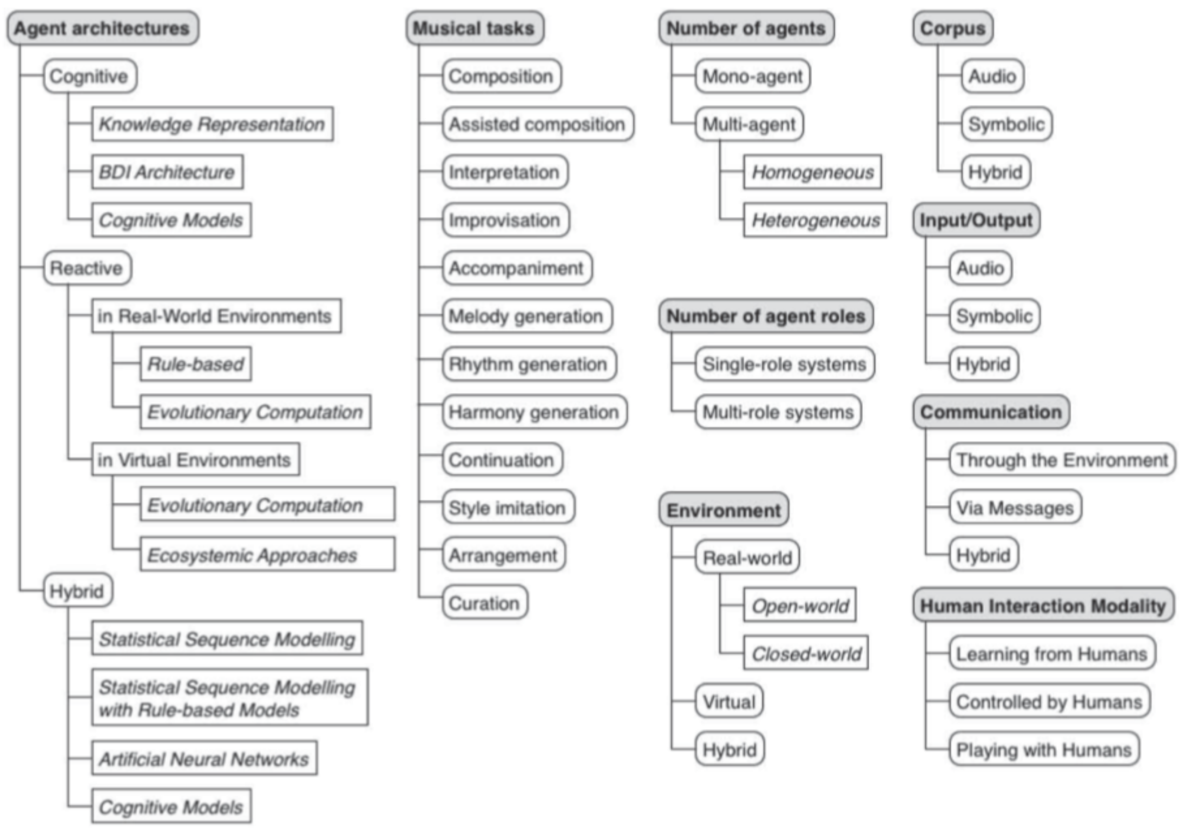


Figure 1. A typology of musical agents (Tatar and Pasquier 63)

## System Description

*Interdependence*

The behaviour of *Intergestura* takes its structure from Oliveros' *Interdependence*, a piece from her collection *Four Meditations for Orchestra* (Oliveros 1997). There are two roles for performers in *Interdependence* - sender and receiver. The piece directs senders to play a short, staccato pitch and receivers to respond to a short pitch with a short pitch. Performers can switch between these two behaviours at will. There are three variations on this base call and response structure (Table 1). *Intergestura* uses gestures as its fundamental unit. As such, references to 'pitch' or 'tone' in the instructions for *Interdependence* are replaced with 'gesture' in *Intergestura*'s structure. For instance, the base version becomes 'respond to a short gesture with a short gesture.' *Interdependence* is an exercise in careful, attentive listening and instantaneous response. Oliveros notes that "correct player reactions can create an atmosphere of electricity that runs through the ensemble in a rippling effect."

| **Variation** | **Respond To** | **Respond With** |
|---:|---|---|
| 0 | Short pitch | Short pitch |
| 1 | Short pitch | Short pitch or long tone |
| 2 | Short pitch or end of long tone | Short pitch or long tone |
| 3 | Short pitch or end of long tone | Shirt pitch, long tone, or long tone with glissando |

Table 1. The four variations for receiver behaviour in *Interdependence.* Sender behaviour is always to play a short pitch, at any time and at any dynamic.

The *Intergestura* system (Figure 2) is built in the Max/MSP programming environment. It is comprised of a human performer who interacts gesturally with the system, a pair of granular synthesis modules, memory modules, an agent that plays one of the synthesis modules according to the instructions of *Interdependence*, and a set of audio transformation processes. It features a single agent, making it a *mono-agent* system. It is a *multi-role* system, meaning that the agent can act as either sender or receiver as described. In terms of task, the agent both *interprets* the instructions of *Interdependence* and *improvises*, in a constrained fashion, with the human performer.

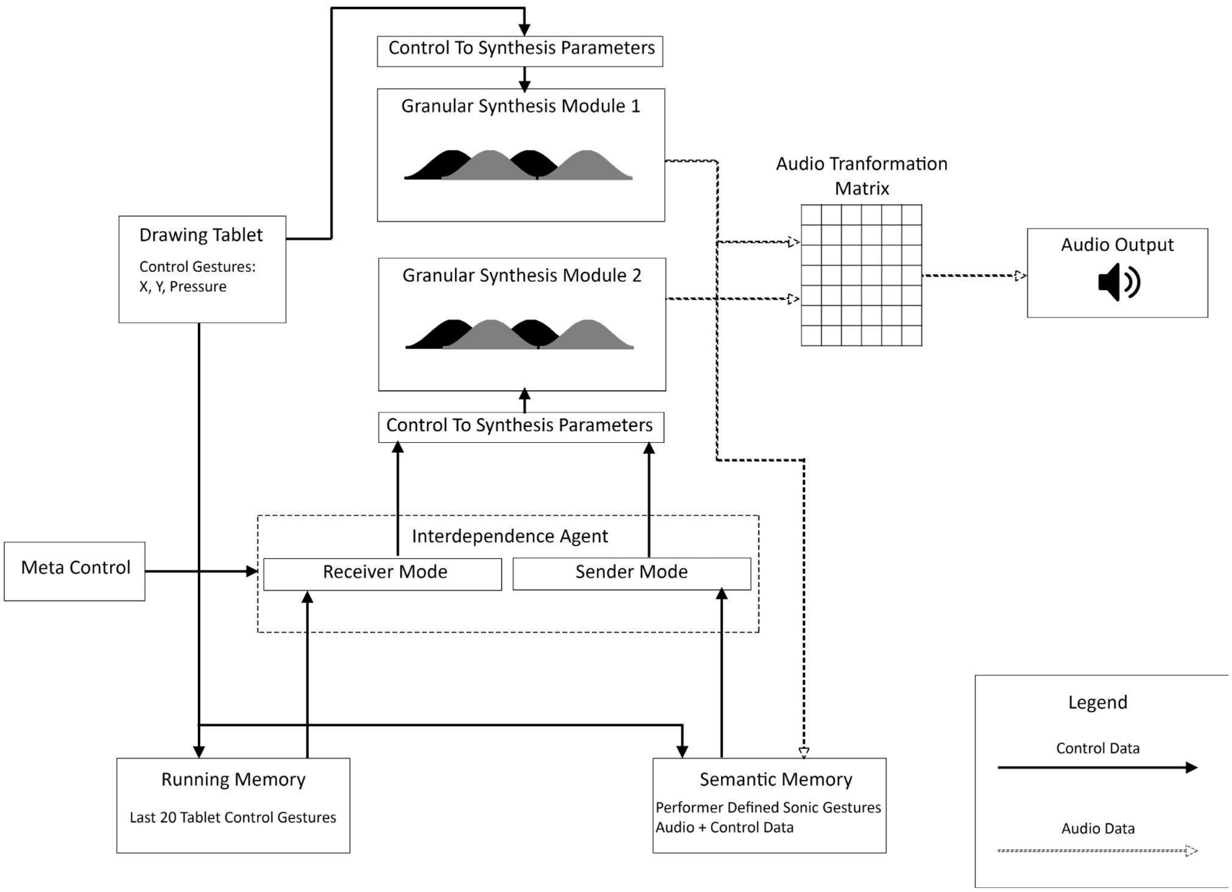


Figure 2. *Interdependence Gestural Agent* system diagram

*Gesture*

Gesture plays a fundamental role in the system. Human input into the system primarily takes place through gestural input with a digital drawing tablet. Physical gestures on the tablet are turned into control gestures and through various mapping processes connected to one of the granular synthesis modules, finally producing an audible sonic gesture (Van Nort 178). Additional control into the system takes place through a standard MIDI controller with banks of knobs, sliders, and buttons.

Gesture is one way that meaning is produced and understood in music-making (Leman 126). Through physical and sonic gesture, meaning is made directly and indirectly - Leman writes "gesture appears as a mediator for music-driven social interaction or as the vehicle through which a 'me-to-you' relationship is established" (142). Here Leman is referring to the relationship between a piece of music and someone experiencing it, but this view on gesture is also a potentially rich space for thinking about the development of agential systems. In *Intergestura,* an important part of agent behaviour is based on call and response, listening and reacting. Human gesture is how the agent understands its role, and it is through gesture that the human-machine relationship is established. Citing Broeckx, (1981) Leman additionally writes, "through gesture, music can be experienced as the action of a dynamic organism similar to a human organism" (126). By having agent interaction take place gesturally, the agent becomes part of this dynamic organism and can participate in the co-creative process.

### *Granular Synthesis Modules*

Sound in *Intergestura* is produced through a pair of identical granular synthesis modules. The human performer plays one module and the agent plays the other. Each module contains a pair of granular engines that run in parallel to each other. Each engine receives the exact same set of synthesis parameters - the same sound source, grain size, pitch, volume, etc. Each engine provides a different quality of sonic granulation - one produces smooth, synchronous, overlapping grains and the other can produce variably spaced, separable grains. There are a variety of textures that can be produced by the two engines, and the performer can crossfade between engines or have them both sounding simultaneously.

Sound sources for granulation are grouped into folders or sound sets, with individual sound sets containing groups of sound sources. Sound sets are generally comprised of three to twelve individual sound sources. Sources are generally grouped with similar sounds in a sound set, though some sound sets contain contrasting sources. For example, one sound set is thematically organized, with all sounds coming from an acoustic guitar - the set contains both harmonic, tonal sources as well as noisier material like the scraping of the strings. In performance, navigation between different sound sets is done through a simple mapping to a single MIDI knob. Navigation within a sound set happens in one of two ways - pressing a MIDI button will randomly select a source from the currently selected set, or navigation can take place through the drawing tablet's touch detection. This means that the sources within a set are laid out on the surface of the tablet and can be moved between by placing and moving a finger across the tablet.

Gestural control data from stylus actions on the tablet are mapped into the granular module to produce sound. There are a number of mapping modes that can be selected. Stylus pressure is always mapped to the volume of the synthesizer output. In the most direct mode, stylus y-position is mapped to both grain rate (how often a grain 'fires') and grain size. X-position is mapped to scrub position in the selected sound source. Another mode makes use of a self-organizing map (SOM) neural network. In this SOM mapping, granular synthesis parameters are two-dimensionally organized across the surface of the tablet and are navigated through according to stylus coordinates. Scrubbing through the sound source is done through a time-based method inspired by Van Nort et al. (2014, 12). Lastly, there is a mapping mode that makes use of a multilayer perceptron (MLP). Inputs into the MLP are the stylus coordinates as well as the stylus velocity, acceleration, and jerk magnitudes. Outputs from the MLP are all granular synthesis parameters. Faster, more chaotic gestures produce noisier sonic output while slower, smoother gestures produce smoother sonic output. In all mapping modes, the final sonic output is highly dependent upon the selected source material.

Finally, the system is greatly inspired by and owes a debt to Doug Van Nort's *Greis* system (Van Nort et al. 2013, 308). Elements such as tablet-based interaction, parallel granular engines, and the structures of semantic and running memories have been influenced by *Greis*.

*Transformation Matrix*

The next component to *Intergestura* is the transformation matrix. This is a collection of audio processes through which the two granular modules can be routed. The processes include distortion, spectral stretching, reverb, and delays. The outputs of these processes can also be routed to each other through the matrix. Routings of the matrix are set by the human performer.

The most important of these processes, as will be discussed later, is the variable-rate delay process. This is a delay process with 8 pairs of delay taps. When the delay time for a pair of taps is changed, the pair crossfades between the two taps, resulting in a changing delay time without clicks or pitch changes. Each of the granular modules processes 8-channels of grains and each of these channels can be sent to one of the delay pairs. Depending on the delay times, this results in modified sonic gestures that feel like they are collapsing or being stretched out or broken apart.

The delay times and feedback levels can be set interactively. There is a machine listening process that calculates the yin, generally of the output of the human-controlled granular module. The 4-dimensional output of the yin process is used as the input to an MLP. The output of the MLP is a vector that sets the time for each delay line as well as its corresponding feedback level. The MLP is trained on different sonic states with different sets of delay times and feedback levels for each state. For example, a high pitched harmonic sound from the granular module will result in one set of delay times that stretches out a gesture, while mid-range noisy content from the granular module will result in a different set of delay times that compress a gesture.

*Agent Module*

*Intergestura* has a *hybrid* architecture. It uses the cognitive concepts of semantic and episodic memory to form the agent's corpus in combination with a reactive rule-based system, inheriting the instructions from the *Interdependence* meditation. Each form of memory, semantic and episodic, can be considered to be constructive of its own corpus. This is one dimension of the typology that is complicated by the system's gestural nature. The semantic memory constructs a corpus that is *hybrid*, containing control gesture, sonic gesture, and machine analysis of the gesture. The episodic, or running, memory constructs a corpus that is *gestural*, consisting only of control gesture data. This also speaks to the agent's environment type - it is similarly complicated in that "real-world" environments in Tatar and Pasquier's typology consist only of audio. *Intergestura's* environment is *real-world*, but adds gestural data to the environment and additionally limits the moments when audio data is available to the agent.

Audio is available to the agent only when it is acting in sender mode and making use of semantic memories. Semantic memories are sonic gestures that are explicitly segmented by the human performer. This is done in performance by pressing a button on the stylus at the start of the gesture and again at the end of the gesture. The gesture, containing control data and audio, is then added to the semantic memory and analysis performed on the audio component of the gesture. There are up to 10 semantic memories at any given point, with the most recent memory taking the place of the oldest memory.

The running memory does not store sonic gestures, only control gestures, which are used by the agent when it is in receiver mode. The running memory consists of the last 20 gestures that have been made by the human performer. These gestures are segmented automatically - segmentation starts when the stylus makes contact with the tablet and ends when the stylus is lifted. Gesture length is calculated and stored on segmentation end. Again, incoming memories replace the oldest stored memory.

At the beginning of sender mode behaviour, the agent randomly selects one of the gestures from the bank of semantic memories. This gesture is used to structure the agent's behaviour. It then selects a number of short gestures to play into its corresponding granular module. The agent looks at the rate of onsets detected in the analysis of the sonic component of the gesture and uses this rate to determine how often it should play its short gestures. Each time it wants to make a short gesture, it randomly selects a segment from the recalled memory. If the segment is above a certain threshold in length, it will be temporally compressed. The control data from the segment is then played into the granular module at the determined rate. After the predetermined number of segments have been played back, the agent then makes a choice on whether to choose a new semantic memory and repeat the process, or to enter receiver mode. The probability of this choice is human-controlled.

On entering receiver mode, the agent randomly selects a number of gestures to respond to. It then starts to look for the end of human-input gestures, specifically waiting for a message that the stylus has stopped making contact with the tablet - this makes the agent's communication type *message-based*. On lift of the stylus, the agent randomly selects a gesture from the running memory. How it treats this gesture differs based on which variation it is performing. In variations 0 and 1, the agent checks the length of the last input gesture. If the length is above a threshold, the agent does not respond and returns to listening for the next stylus lift. If the length of the input gesture is below the threshold, the agent looks at the length of the recalled gesture. In variation 0, the agent can only respond with short gestures. If the length of the recalled gesture is above a threshold, it temporally compresses the gesture and plays the control data into the granular module. If the gesture is below the threshold, it simply plays it back without any compression. After the gesture is played back, the agent continues to listen for stylus lifts - note though that if a stylus lift is detected while a short gesture is being played, that gesture can be interrupted and a new gesture selected and played back. When the agent has played back its predetermined number of gestures, selected at the start of the process, it then has the option to remain in receiver mode or return to sender mode. The probability of this is, again, human set. If the agent chooses to remain in receiver mode, it additionally randomly selects which variation it will perform and returns to the top of the process.

Each of the variations adds additional logic for how recalled gestures are handled. Variation 1 gives the agent the ability to respond with long gestures instead of solely with short ones. To achieve this, after it has recalled a gesture for playback, the agent again checks its length. In this case, if the length is *below* the threshold it is compressed slightly, making short gestures shorter. If it is *above* the threshold, it stretches the gesture. The closer the length is to the threshold, the more it will be compressed or stretched. This behaviour pushes the lengths of played back gestures to the extremes, with more short and long gestures and fewer gestures of medium duration. Variations 2 and 3 eschew checking the length of the most recent input gesture - the agent responds to gestures of any length. The third variation gives the agent the option to add random, narrow-interval glissandi to the played back gestures.

Through the sender and receiver behaviours, the I/O type of the system can be determined to be a *hybrid*, with both audio and gestural data being used for input and gestural data being used to generate audio for output.

As mentioned, there are controls that the human performer sets for the agent - making the system's Human Interaction Modality one of *systems controlled by humans.* These controls can be thought of as meta-controls - they do not directly set or control the agent's behaviour, but influence the probability that the agent will behave in a certain way, additionally acting as a sort of stability control. There are two such controls, each mapped to a physical MIDI knob and slider, respectively. The first meta-control influences the likelihood of the agent staying or switching between sender and receiver modes. At either extreme, it will remain in the respective mode until the control is changed. The second meta-control maps to parameters that exclusively influence the behaviour of the agent in receiver mode. Towards one end of the control, the agent will be more likely to play variations 0 or 1 and towards the other end, it will be more likely to play variations 2 or 3. At the center point, it is equally likely to select from any of the variations. At either extreme of the control, the agent will stop selecting new gestures for playback and continuously repeat the last selected gesture.

Finally, the agent also has some autonomy over the sound source selected in the granular module. While it is constrained to the sound set selected by the human performer, every time the agent receives an input gesture, it also looks at the standard deviation of the velocity of that input gesture. If it is above a threshold, the agent will randomly select one of the sound sources from the given sound set for granulation.

**Discussion**

Playing with *Intergestura* certainly feels like a metacreative process of human-machine co-creation. The agent is highly responsive and it often feels like interacting with a mirrored version of myself - it is always interesting to hear when the agent returns a gesture that is recognizable as something that I recently played, but has been transformed temporally or through a different sound source. Both Brown and Ravikumar et al. argue for the creation of human-machine systems that embody co-creativity through partnership and shared tasks. Brown says, “we should look beyond metaphors of dialogue or interaction and be concerned with partnerships with generative processes where the ‘system’ includes a human participant, as opposed to a system as a tool or an autonomous machine” (2012, 18). Ravikumar et al. additionally argue that co-creation should happen through a process of co-experience, “the experience of musical co-creation that is co-improvised as opposed to improvised alone… co-experience is meant to be an experiential construct that measures aspects of feeling together during an interaction” (2018, 3).

Attempting to perform *Interdependence* with this system seems to embody this definition of co-experience. As with performing any *Sonic Meditation*, I find it requires attentive listening and quick responses. I never really know exactly what the system is going to throw at me, whether it is going to respond to me or act as a sender. This means that I must be ready to respond at a moment's notice, just as in improvising with human players. As *Sonic Meditations* blur the line between performer and audience member, I similarly find that performing with *Intergestura* can blur the line between myself and the agent. When we get into a back and forth of many quick gestures in succession, it can be difficult to tell which gestures I produced and which came as reflections from the system. In listening back to recordings of performances, this becomes even more difficult. Furthermore, playing with the system but *not* trying to play *Interdependence* is also an interesting experience. In this type of playing, the agent feels more like an augmenter of my own performance, allowing me to create sounds and shapes that I would not be able to do on my own.

The quick back and forth style of playing described above can result in feeling like the ripples of electricity that Oliveros describes, but when longer gestures are introduced this feeling often is not present. One way I have found of creating this rippling feeling with just myself and the single agent is to route either or both of the granular modules to the variable delay process. This machine-augmentation for both myself and the agent results in another layer of interaction and can result in densely rippling textures. The variable delay lines have, in fact, become an indispensable part of performing with *Intergestura*.

While it can be and often is a very satisfying experience, the system can at times feel obstinate. While it is desirable for the agent to have agency and be able to lead the co-creation process, I don't feel that it is always successful at this task. In particular, the sender mode of the agent often leaves something to be desired. For example, sometimes when switching from receiver to sender mode, it can feel as if the agent has decided to ignore the musical direction the performance was taking. This *can* be interesting, and when it is it can feel magical, but often it is simply frustrating and in response to this I find myself turning up the meta-control to keep the agent mostly in receiver mode. I believe that a redesign of the sender behaviour that can take into account the current sonic direction would greatly improve the agent.

Additionally, the recalled gestures do not always feel appropriate for the sound content selected in the granular module. In my own playing, I often find that the sound source, especially whether it is a tonal or noisy source, dictates the type of gestures that I might make. The agent, as is, doesn't have any way to take into account the quality of the sound source it is working with. I believe that building in classification of sound sources, giving the agent knowledge of these classifications, and giving it the ability to select different mapping modes based on this information would be one way to improve its behaviour.

When I first began work on this system I had conceived of it as a digital musical instrument meant for improvisational playing. But as I have developed and played with the system I've found this conception doesn't seem to work. As other lines have been blurred, *Intergestura* itself lies somewhere between instrument, improvisational system, and composition. Arne Eigenfeldt's notion of real-time composition (RTC) systems perhaps best describes *Intergestura*. "Real-time composition is the application of musical agents to interact in musical ways, during performance… knowledge of how to interact musically must be built into the agents, and an environment created in which these agent interactions can result in artistically interesting and compositionally satisfying sonic artworks" (Eigenfeldt 146). RTC systems, he writes, are designed as a singular composition that can change drastically from performance to performance. With its designed agent interactions and ability to produce radically different outputs in different performances, *Intergestura* seems to fall well into this definition, thus making the system a real-time composition and co-creative partnership.

**Future Work**

In addition to the improvements mentioned in the previous section, I believe one interesting direction to take *Intergestura* would simply be to add more pairs of agents and granular synthesis modules. In this multi-agent formation, it would also be necessary to give each of the agents awareness of each other in addition to the human performer. By doing so, the agents would be able to respond to gestures from each other. There would need to be built into the agents a notion of attention and awareness - is a given agent listening to other agents, the human performer, or the entire group? I imagine with even just one additional agent this could result in long chains of rippling actions and reactions.

An alternate version of this system could be one in which all interaction takes place through sound. As in *dispersion.eLabOrate*, this version of the system could be used to augment traditional *Sonic Meditation* practices. I imagine it could be able to perform within an ensemble of humans with diverse instrumentation and engage in co-creation purely through sound. This would require a significantly different approach than the current gestural approach but would be great in opening up interaction with the system to a wider range of people.

## References


Broeckx, Jan L. *Muziek, Ratio En Affect: Over De Wisselwerking Van Rationeel Denken En Affectief Beleven Bij Voortbrengst En Ontvangst Van Muziek*. Metropolis, 1981.

Brown, Andrew R. "Eighth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment." *1st International Workshop on Musical Metacreation* , 2012, https://musicalmetacreation.org/proceedings/mume-2012/. Accessed 21 Dec. 2021.

Brown, Andrew R. "Ripples: A Human-Machine Musical Duet." *The Fourth International Workshop on Musical Metacreation*, 2016, https://musicalmetacreation.org/proceedings/mume-2016/. Accessed 21 Dec. 2021.

De Souza, Jonathan. "Orchestra Machines, Old and New." *Organised Sound*, vol. 23, no. 2, 2018, pp. 156–166., https://doi.org/10.1017/s1355771818000031.

Eigenfeldt, Arne. "Real-Time Composition as Performance Ecosystem." *Organised Sound*, vol. 16, no. 2, 2011, pp. 145–153., https://doi.org/10.1017/s1355771811000094.

Gioti, Artemi-Maria. "From Artificial to Extended Intelligence in Music Composition." *Organised Sound*, vol. 25, no. 1, 2020, pp. 25–32., https://doi.org/10.1017/s1355771819000438.

Hoy, Rory, and Doug Van Nort. "An Ecosystemic Approach to Augmenting Sonic Meditation Practices." *14th International Symposium on Computer Music Multidisciplinary Research*, 2019, pp. 318–325.

Leman, Marc. "Music, Gesture, and the Formation of Embodied Meaning." *Musical Gestures: Sound, Movement, and Meaning*, edited by Rolf Inge Godøy and Marc Leman, Routledge, New York, 2010, pp. 126–153.

Lewis, George E. "Too Many Notes: Computers, Complexity and Culture in Voyager." *Leonardo Music Journal*, vol. 10, 2000, pp. 33–39., https://doi.org/10.1162/096112100570585.

Magnusson, Thor. *Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions*. Bloomsbury Academic, 2019.

Oliveros, Pauline. *Software for People: Collected Writings 1963-80*. Station Hill Press, 1984.

Oliveros, Pauline. *Sonic Meditations: March - November 1971*. Smith Publ. American Music, 1974.

Oliveros, Pauline.  Four Meditations for Orchestra. 1997.

Ravikumar, Prashanth Thattai, et al. “Back to the Experiences: Empirically Grounding the Development of Musical Co-Creative Partners in Co-Experiences.” *6th International Workshop on Musical Metacreation*, 2018, https://musicalmetacreation.org/proceedings/mume-2018/. Accessed 21 Dec. 2021.

Straebel, Volker, and Wilm Thoben. “Alvin Lucier's Music for Solo Performer: Experimental Music beyond Sonification.” *Organised Sound*, vol. 19, no. 1, 2014, pp. 17–29., https://doi.org/10.1017/s135577181300037x.

Tatar, Kıvanç, and Philippe Pasquier. “Musical Agents: A Typology and State of the Art towards Musical Metacreation.” *Journal of New Music Research*, vol. 48, no. 1, 2018, pp. 56–105., https://doi.org/10.1080/09298215.2018.1511736.

Van Nort, Doug, et al. “Electro/Acoustic Improvisation and Deeply Listening Machines.” *Journal of New Music Research*, vol. 42, no. 4, 2013, pp. 303–324., https://doi.org/10.1080/09298215.2013.860465.

Van Nort, Doug, et al. “Mapping Control Structures for Sound Synthesis: Functional and Topological Perspectives.” *Computer Music Journal*, vol. 38, no. 3, 2014, pp. 6–22., https://doi.org/10.1162/comj_a_00253.

Van Nort, Doug. “Instrumental Listening: Sonic Gesture as Design Principle.” *Organised Sound*, vol. 14, no. 2, 2009, pp. 177–187., https://doi.org/10.1017/s1355771809000284.

Whalley, Ian. “Software Agents in Music and Sound Art Research/Creative Work: Current State and a Possible Direction.” *Organised Sound*, vol. 14, no. 2, 2009, pp. 156–167., https://doi.org/10.1017/s1355771809000260.

Wu, Jiayue Cecilia. “From Physical to Spiritual: Defining the Practice of Embodied Sonic Meditation.” *Organised Sound*, vol. 25, no. 3, 2020, pp. 307–320., https://doi.org/10.1017/s1355771820000266.