MULTIMODAL ROBOTIC ARTISTIC INSTALLATIONS FOR SOUND PRODUCTION

Artemis Moroni
Robotics and Computer Vision Division / Center for Information Technology Renato Archer – DRVC/CTI
Artemis.Moroni@cti.gov.br

Jônatas Manzolli
Music Department and Interdisciplinary Nucleus for Sound Studies / University of Campinas – DM/IA, NICS/UNICAMP
jonatas@nics.unicamp.br

Abstract: This article describes sensing systems and mobile robotic platforms that were developed as multimodal robotic artistic installations for sound production. AURAL, an evolutionary mapping of trajectories of the robots into sound events is compared to $AURAL_2$, a generative soundscape described as a regulating system. Aspects of man-machine and machine-machine interaction are approached in perspective to viewpoints related to Computational Creativity and Evolutionary Sonification.

Keywords: evolutionary computation, general system theory, generative art, interactivity, robotics

INTRODUCTION

Throughout Western History, autonomous machines have been often seen connected with music production. In Ancient Greece Ctesibius (c. 270 BC) applied knowledge of pneumatics and hydraulics to produce the first organ and water clock with moving figures. In Middle Ages (c. 1206) it is believed that Al-Jazari, long resident in Turkey, created a programmable set of automata that played music. In 1495 Leonardo da Vinci designed a robot knight that responded to a drum beat (Rosheim, 1994). Despite creation of machines to operate autonomously dates back to classical times, research in functionality and potential use of robots did not grow substantially until the 20th century, but nowadays artists' ever increasing interest in autonomous machines is present in many different contexts, including installations, performances, new instrument designs and collaborations with robotic performers, interactively permeated. When robots, also called real-world devices, are used in an artistic context, it seems important to discuss aspect of creativity and artificial intelligence, as described by Boden (1998). In this way, this article introduces concepts on the computational simulation of creativity in perspective of two systems created by the authors: AURAL and $AURAL_2$.

Focusing on an interaction metaphor that a robotic device is a suitable interface to establish a connection between the virtual and the real world, sensing systems and mobile robotic platforms were developed for AURAL and $AURAL_2$ systems. These multimodal sound environments supply a platform for robotic experimentation and artistic creation exploring the *arTbitrariness*, a framework for developing automatic and semi-automatic processes applied to artistic production, in visual and sound domains (Moroni et al., 2006). arTbitrariness refers to the idea of emulating some aesthetical judgment, arbitrary, through computational techniques. Strongly based on interactivity, arTbitrariness explores human and machine creativity bringing about results that could not be obtained without such interplay.

In AURAL and $AURAL_2$ environments, sonification was generated in two different ways:

1) evolutionary mapping of the trajectories of the robots into sound events (Moroni, 2010) and 2) a generative soundscape characterizing a regulating system. The issue mainly addressed in this article is to compare these two systems concerning to structure/novelty tradeoff such as described by (Todd & Werner, 1999).

The next section introduces main topics for discussion: evolutionary computation as a tool to emulate creativity in computers, followed by the description of AURAL system, where this paradigm is applied. Next, $AURAL_2$ is described, preceding a discussion on automation versus interactivity. Finally, an analysis on the man-machine interactivity modalities that occurred on the two systems is presented, concerning to the sound organizations produced by man-machine and machine-machine interaction.

Creativity, Evolution and Novelty

Recently, evolutionary systems (Bentley, 2001) have been applied to emulate creativity in computers. Among the aspects that justify the use of evolutionary computation techniques, is the fact these algorithms are based on *population search techniques*. Simulated-evolution techniques are useful tools for searching large spaces using operators of variability and selection to obtain "new" material. Indeed, it seems that the common operations and procedures of computational evolutionary processes (crossover, mutation, evaluation, selection, reproduction) appear to be compatible and able to simulate the "*three kinds of creativity*" nominated by Boden (1998): *combinatory*, *exploratory* and *transformational creativity*. The crossover operator strongly contributes to the *combination* of solutions, the mutation operator for the *transformation* and the other procedures for the *exploration* of the domain. But, independently of this, the search algorithms require the definition of an individual evaluation for each solution.

Todd & Werner (1999) discussed the structure/novelty tradeoff in this way: "*more highly structured systems can produce more highly constrained output*". In algorithmic composition systems, this means that more knowledge and structure allows the creation of new compositions that are more tightly matched to the desired musical genre. The flipside of more structure is less novelty: the highly constrained output will be less likely to stray beyond a genre's bound or be surprising. Thus, the highly structure composition system will be less general, able to reach less of "music space" with its output.

AURAL: Evolutionary Sonification

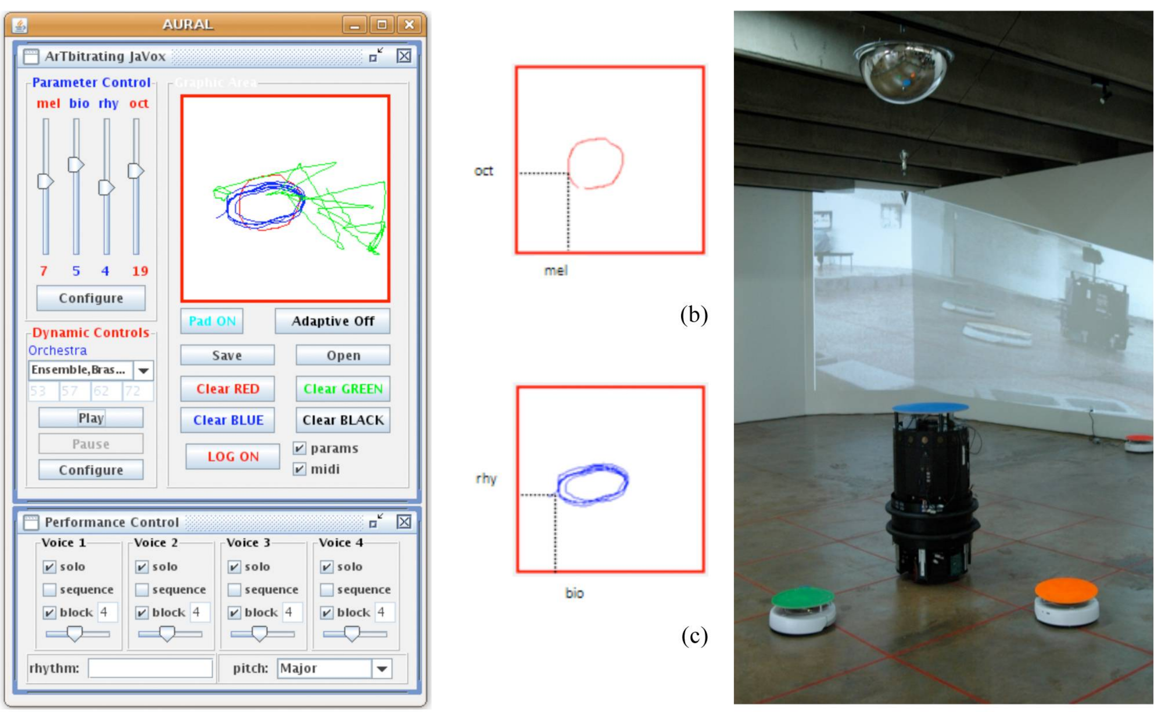
In AURAL, the conflict described in the last paragraph is treated through the interaction among an evolutionary sound process, an artificial vision system and mobile robots. In the sound control interface there is a *Graphic Area*, the heart of the system, wherein the user draws curves that are used to control robotic trajectories. Red curves are sent as trajectories to the robots, and guide the evolutionary sound process across different regions of the sound space. Further, paths traversed by the robots, in an arena, are observed by an artificial vision system that produces sequences of coordinates (x, y), plotted as blue curves in the Graphic Area. This data is fedback into the evolutionary sonification module.

Figure 1 shows, on the left, the AURAL Graphic Area. Three curves are shown: *a)* the trajectory draw by the user, *b)* the path followed by a master robot and *c)* the path followed by another robot. Curves *a)* and *b)* are shown in detail on the middle. Using this evolutionary cycle the initial control input (the red curve) is transformed by the dynamic of the robots in real time. The sound output produced by the AURAL is sensitive to the organization expressed by the behavior of the robots in the arena.

In the evolutionary cycle, the individuals of the population are defined as groups of four notes (Moroni & Manzolli, 2010). Initially, these notes are randomly generated in the interval [0, 127] with each value representing a MIDI event. In each generation, 30 groups are created. Considering that there are 128^4 possible groups of four notes or chords, a population of 30 individuals represents a huge reduction in the dimension of the search space. During several iterations of the system, the initial population evolves to new configurations of chords guided by the fitness evaluation.

The musical fitness of each chord consists of three partial fitness functions: *melodic, harmonic and vocal range*, each resulting in a numerical value (Moroni et al., 2002). The set of notes with the highest fitness is selected and played as a new MIDI event, the duration of the evolutionary cycle (*bio*) and music meter (*rhy*) is taken into account.

The fitness criteria, based on the ordering of consonance of musical intervals, introduces in the process some structure and knowledge. At the same time, depending on the distance between the pairs of robots (until four), performance controls are activated. This strategy introduces sound information of second order and brings about emergent and unexpected output, using data stored in the recent memory of the system. The *Performance Control* area offers other possibilities to control the sound production. For each of the four MIDI events there are three controls: *solo, sequence* and *block*. They work as delay lines in which MIDI Note events from previous generations are played again as solo, melodic patterns or chords. These controls are also modified in real time by the behavior of the robots. They are used to select the solo, the sequence or block mode for each voice. Table 1 shows the five simple rules associating the distance between the robots and processes (the solo, sequence and block) of the Performance Control.



(a) (d)

Fig. 1 - On the left (a), the AURAL interface shows the different control areas: the Parameter Control; the Graphic Area and the Performance Control. On the middle, there are details of the curves in the Graphic Area. On the top (b), the parameters extracted from the trajectory that was sent to a master robot. On the bottom (c), the path followed by the robot. On the right (d), there is an image of AURAL as an art installation. Hanging on the ceiling, the artificial vision system, specially designed for the robot tracking.

Other interface feature enables the user to modify number of notes, rhythmic pattern, basic pitch reference and orchestra (General MIDI instruments), which also affect the musical performance. The user interaction can be interpreted as attempts to improve the outcome, opening the possibility of the system to learn with it. During an AURAL performance, all the interactive paths can be recorded. It is possible to record all the automatic and interactive events, as well as the audio and MIDI files generated in real time. Some of them were used as a basic material for generating instrumental compositions. A composition titled “Robotic Variations” (Jornal Nacional, 2009; AURAL Robotic Sonification, 2011) for Piano, Marimba and Electronics (computer and robots) was composed from the obtained sound material and performed at the AURAL installation.

Table 1. Rules relating the distance between the robots and Performance Controls

Rule	Distance (m)	Solo	Sequence	Block
1	>0.5	X		
2	0.4 < D < 0.5		X	
3	0.2 < D < 0.4		X	X
4	D < 0.2			X

AURAL Installation

AURAL was presented in an Art Gallery in March 2009 (Figure 1.d) where the visitors could appreciate the sound output and the interaction among the robots, as a kind of choreography. The visitors drew curves in the JaVOX GUI, which were transmitted as trajectories to the master robot, Nomad. While the robots (until 4) moved in the arena, virtually traveling along the conceptual sound space, people changed the orchestra, rhythm and pitch controls, investigating the sound possibilities. Both a process of man-machine interaction and parallel exploration occurred. On the last day of the exhibition, a dancer, Tatiana Benone, three musicians, Cesar Traldi, Adriano Monteiro, Francisco Costa and the AURAL system itself, with four robots, performed the interactive concert “Robotic Variations” (Jornal Nacional, 2009; AURAL Robotic Sonification, 2011). The same trajectories used to generate the material for the composition were used in the performance. For the visual tracking, a strong color panel is fixed on the top of each robot. An interactive scenario displayed real time processed images on the walls. Figure 2 shows images from the performance. Video material is available at (AURAL Robotic Sonification, 2011).

The dancer was invited to interact with the robots in the arena, in a live performance. Choreography was designed so that the Create robot with a red panel left the room and was substituted by the dancer using a red hat. Her position was tracked by the visual system through the red hat and interfered in the performance of the sound, incurring in another human-machine interaction cycle. Figure 3 shows some pictures of the musicians and of the dancer taken during the rehearsals for the performance. This performance exploits a conceptual sound space using robots and human agency, in which the musicians and the dancer used pre-composed material and improvisation. The evolutionary system, even when performed using previous control parameters, produced new sound events differing from the previous recorded MIDI files. Nevertheless, there was an overall organization reflecting the collaboration among the agents.

AURAL$_2$: Generative Sonification

A similar architecture, with an artificial vision system and mobile robots but with a different sonification paradigm was applied in AURAL$_2$. If in the previous AURAL, the sound production resulted from an evolutionary process, in the second version the result came from a generative process. Generative systems have many similarities with systems found in various areas of science. They can simulate behavior related to order and disorder, as well as a varying degree of complexity, making long-term prediction difficult. However, such systems still contain a definite relation between cause and effect. The artist (or designer) generally provides basic rules, and then defines a process, random or semi-random, to organize these elements. The results continue to happen within the limits of the rules domain, but also may be subject to subtle changes or even surprising.

Fig. 2 - The pictures above shows the dancer, the scenery, the omnidirectional vision system (hanging on the ceiling), the robots and the musicians, during the rehearsals for the art performance.

$AURAL_2$ was designed to generate a soundscape associating virtual and real world to sound production. First off all three elements were related: a) a sound database with synthetic, game and environmental sound samples - the sound memory of the system, b) a computer graphic grid - the virtual world and c) a winding platform (3m x 3m wide, 0.3m high) - the stage (Figure 3). The following association correlated these three elements: "*each sound within the database was associated with a cell in the graphic grid*".

The position of the robots on the platform is mapped to a cell, through a webcam. Robot trajectories across different regions of the stage trigger sounds associated to cells of the grid, (re)creating a soundscape in the installation environment. Secondly the borders of the stage are irregular, and there is a round hole in the middle of this platform. This feature creates two regions on the stage: a) one region can be traversed only by one robot and b) the other region by one or two robots.

The robots have an edge sensor and they turn over when detect an edge. In other regions of the platform, three or four robots can move freely. This design sometimes promotes spatial conflict among the robots while they are trying to escape from a confined area. Solutions for the conflict produce unexpected sound combinations – novelty that can be associated to the concept of combinatory creativity, mentioned above (Boden, 1998).

On a TV display, the virtual grid is shown in various angles, as well as the cells activated by the trajectories of the robots. Visitors can also interact with $AURAL_2$ by speaking, singing or screaming at a microphone. Sound fragments are extracted from the visitor's interventions and are randomly inserted into the database. Further, the movement of the robots can trigger these sound fragments again.

A spectral analysis is also applied on the fragments of the interventions to generate two visual effects: a) when there is more energy in upper partials, the color of the cell associated with that fragment is changed to reddish aspect, otherwise to bluish. A rotation is applied on the grid; the sound fragment is inserted in a random position of the sound database, deleting a previous sound, in an acoustical recycling process. This generative cycle of transformations produced by $AURAL_2$ can be seemed as "sound ecology". Therefore $AURAL_2$ intervenes in the acoustic environment generating new aural trajectories with everyday sounds.

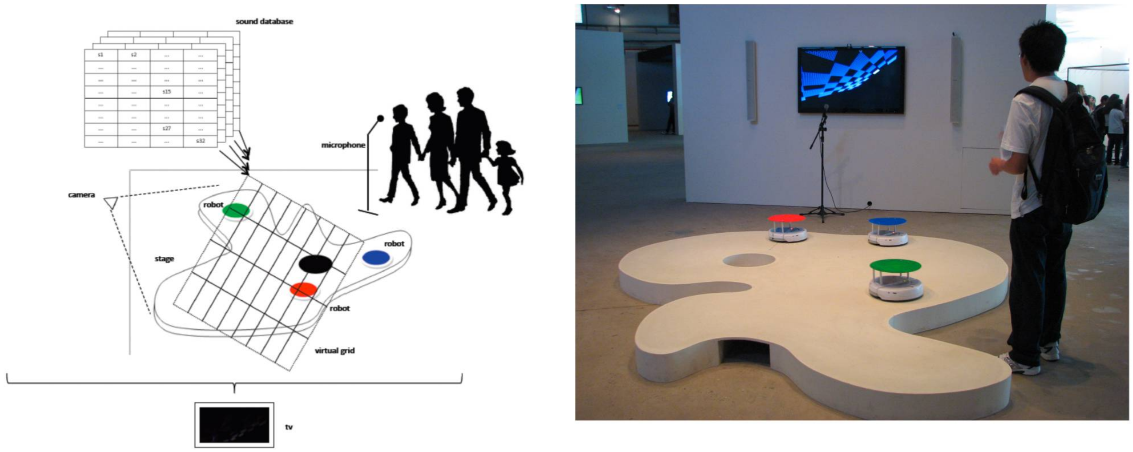


Fig. 3 – On the left, there is a diagram of AURAL2 installation. A virtual grid associates different sound databases with the platform: synthetic sounds, game sounds, everyday sounds and environment sounds. The movement of each robot - its location, monitored by the camera - triggers the sound associated with that place in the grid. On the right, there is a picture of the installation.

Automation versus Interactivity

One can see an interesting aspect of AURAL is the possibility of using different setups to explore distinct levels of interaction among humans and machines. One way to characterize types of interactions is to analyze the ways in which systems can be coupled together to interact. Cornock & Edmonds (1973) early identified the concept of ”*Art System*” as consisting of the *artist*, the *participants*, the *artwork*, the *environment* in which these elements are placed, and the *dynamic processes* or *interactions* that result (Candy & Edmonds, 2012). Canonical models of computer-human interaction are based on an archetypal structure: the feedback loop. Representing interaction between a person and a dynamic system as a simple feedback loop is a good first approximation, it forefronts the role of information looping through both person and system (Dubberly et al., 2009).

In AURAL, the user supplies parameters for fitness evaluation by drawing a red curve. The coordinates (x,y) of this curve provided input for an evolutionary process. Medium solutions are expected in this case, since the fitness function changes quickly. The blue curve - output - supplies the *bio* parameter for the reproduction cycle and a rhythmic parameter for the MIDI event cycle. If in AURAL the process is of a *reinforcing* system; in $AURAL_2$, there is a *balancing* system. And, not mentioned but latent, once the better individual of the population is selected (reproduction cycle) and put in a critical area to be played as MIDI event, the process is of a *conversing* system, when the output of a learning system becomes input for another.

In $AURAL_2$, depending on the position of the robots on the stage, sound fragments are triggered, characterizing a *regulating system:* the output of one linear system provides input to another. But, most important about $AURAL_2$, is that it is an *open system*. $AURAL_2$ is sensitive both to the sounds of the environment and the interactions of the visitors at the microphone, storing sound fragments in the database when the result of the spectral analysis surpass a threshold. These fragments can be played again, in a continuous acoustic recycling process, once the cells where they are stored in the virtual grid are activated by the movement of the robots.

Interactive Performance

In 1981 Chadabe (1997, 2005) proposed the term "*interactive composition*" to describe "*a performance process wherein a performer shares control of the music by interacting with a musical instrument*". Programmable interactive computer music systems such as these challenge the traditional clearly delineated western art-music roles of instrument, composer and performer. In interactive music systems the performer can influence, affect and alter the underlying compositional structures, the instrument can take on performer like qualities, and the evolution of the instrument itself may form the basis of a composition. In all cases the composition itself is realized through the process of interaction between performer and instrument, or machine and machine.

Several interactive processes were observed in the AURAL environment. In the interactive concert "Robotic Variations" (Jornal Nacional, 2009; AURAL Robotic Sonification, 2011), for example, the musicians played a music for which the movement of the robots on the arena was used as a composition strategic. A trajectory was sent to Nomad, the master robot that tried to follow it, while other robots navigated on the arena, controlled by a pre-programmed autonomous mode. The same trajectory used to generate the material for the composition was used in the performance, but because of the evolutionary sonification process, even if the parameters of control are alike in every execution, the result is different in each run. The musicians knew the type of music that would be generated, but they had to be able to adapt their performance.

At the same time, the dancer, tracked by her red hat, was interacting with the robots, all interfering in the music that was being generated. In each performance, the place of the robots, navigating in the arena in their autonomous mode, can be different. The dancer had to be able to accomplish them. Important to remember that all the process is triggered by a curve drawn by a human. The same curve was used in the rehearsals and in the final performance, with the musicians, the dancer and the robots. The dancer and the robots interfered in the sonification process, accomplished by the musicians, incurring in multiple feedback cycles.

On the other hand, in the $AURAL_2$ the microphone acted as an invitation to the visitors for interaction. Initially, the people experienced the installation by talking at the microphone. When they hear segments of their speech mixed with other sounds, the visitors started to explore the system by talking, singing, or even screaming. These interventions sometimes incurred in visual effects in the virtual grid displayed on the TV, by changing the color of the cells or the position of the grid. Filtered images of the robots and of the people were also displayed on the TV. When the visitors were aware of the images, they started to move in front of the camera. Their behavior changed while they experienced the environment.

Conclusion

From the AURAL and $AURAL_2$ perspectives, humans and robots are agents of a complex system and the sonification is the emergent propriety that is produced by their interaction and behaviour. This exploration is also related with the concept of self-organization in complex systems. As such the sonification here is not seen as an isolated aspect of these two systems but a representation of the synergetic capacity of the agents to collaborate and produce a complex product. More structure and knowledge built into the system means more reasonably structured musical output, but also more predictable output, which can be relaxed by introducing processes such as those linking the interaction of the robots with the performance controls. Less structure and knowledge in the system, like in $AURAL_2$, means more novel, unexpected output, but also more unstructured musical chaff. The cost of introducing more structure into the system is one of the concerns of the arTbitrariness. Producing computational models of such high-level behaviours, embedded in robotic platforms, calls for novel research at the frontier between robotics, music and multimodal systems.

Acknowledgements

We thank the PIBIC/CNPq Program, the DRVC/CTI and NICS/UNICAMP for making this research possible. AURAL Project was supported by the FAPESP, process 05/56186-9. Manzolli, J. is supported by CNPq.

References

AURAL Robotic Sonification. https://sites.google.com/site/auralroboticsonification/ 2011. Accessed in: 09/06/2012.

BENTLEY, P., CORNE, D. **Creative Evolutionary Systems.** San Francisco: Morgan Kaufmann, 2002.

BODEN, M. "Creativity and artificial intelligence", In: Elsevier Science: Artificial Intelligence, 103 pp. 347 – 356, 1998.

CANDY, L., EDMONDS, E. **Interaction in Art and Technology.** Available at: http://crossings.tcd.ie/issues/2.1/Candy/ Accessed in: 03/06/2012.

CHADABE, J. **Electric Sound: The Past and Promise of Electronic Music.** Upper Saddle River, NJ: Prentice Hall, 1997.

CHADABE, J. "The Meaning of Interaction". **Proceedings of the 2005 HCSNet Conference.** Macquarie University, Sydney, Australia, 2005.

CORNOCK, S. AND EDMONDS, E. "The Creative Process where the Artist is Amplified or Superseded by the Computer". **Leonardo** 6: 11-16, 1973.

DUBBERLY, H., HAUE, U., PANGARO, P. **What is interaction? Are there different types?** 2009. http://www.dubberly.com/articles/what-is-interaction.html Accessed in: 03/06/2012.

Jornal Nacional. http://jornalnacional.globo.com/Telejornais/JN/0,,MUL1052269-10406,00-ROBOS+COMPOEM+MUSICA+PARA+O+HOMEM+NA+UNICAMP.html, 2009. Accessed in: 09/06/2012.

MORONI, A., MANZOLLI, J. "From Evolutionary Composition to Robotic Sonification". In: **EvoApplications 2010.** Applications of Evolutionary Computation. Berlin: Springer, 2010.

MORONI, A., MANZOLLI, J., VON ZUBEN, F. "ArTbitrating JaVOX: Evolution Applied to Visual and Sound Composition". In: **Ibero-American Symposium in Computer Graphics**, 2006.

MORONI, A., MANZOLLI, J., VON ZUBEN, F. J., GUDWIN, R. "Vox Populi: Evolutionary Computation for Music Evolution". In: BENTLEY, P., CORNE, D. **Creative Evolutionary Systems**, pp. 205–221. Morgan Kaufmann, San Francisco, 2002.

ROSHEIM, M. **Robot evolution: the development of anthrobotics**. John Wiley & Sons, 1994.

TODD, P., WERNER, G. "Frankensteinian Methods for Evolutionary Music Composition". In: GRIFFITH N., TODD P. **Musical Networks: Parallel Distributed Perception and Performance.** Cambridge, The MIT Press, 1999.