Dynamic mapping strategies for interactive art installations: an embodied combined HCI HRI HHI approach

Pieter Coussement, Marc Leman, and Michiel Demey
IPEM, Dept. of Musicology Ghent University, Blandijnberg 2, B-9000 Ghent, Belgium
pieter.coussement@ugent.be, marc.leman@ugent.be, michiel.demey@ugent.be

Abstract — This paper proposes a theoretical framework for dealing with the paradigm of interactivity in new media art, and how the broad use of the term in different research fields can lead to some misunderstandings. The paper addresses a conceptual view on how we can implement interaction in new media art from an embodied approach that unites views from HCI, HRI and HHI. The focus is on an intuitive mapping of a multitude of sensor data and to extend upon this using the paradigm of (1) finite state machines (FSM) to address dynamic mapping strategies, (2) medially to address aisthesis and (3) embodiment to address valid mapping strategies originated from natural body movements. The theory put forward is illustrated by a case study.

Index Terms — Art, Artificial Intelligence, Finite state machines, Fuzzy systems, Interactive systems.

I. Introduction

When developing interactive installations we have to be aware of some of the paradigms surrounding interactivity. Addressing interactivity can be done from many different research fields, all posing very different research questions, ranging from psychology to computer science.

A. Social Interaction

As humans, we like to think that interaction is a given, and to some extent this is true; all of us interact with each other on an almost continuous basis, wether we are sitting next to one an other on a bench, are having lunch together or are talking to each other. Apart from our Human Human Interaction (HHI), we spend a great deal of time ‘interacting’ with the world around us in various ways through a multitude of interfaces. These forms of interaction are subject to various research fields that proof to be even more intriguing.

B. Computer Science

In computer science interaction is by far one of the most fashionable words to date, and is associated with research fields as Human Computer Interaction (HCI), Human Robotic Interaction (HRI), among others. How we perceive interaction is very much dependent on our willingness to conceive something as interactive, the context of our interaction and the partner at hand when interacting. Furthermore, ‘being interactive’ is a quality that is often stowed upon an object in a very dubious matter. In such a way that it is often a quality that is perceived by the user, not one that is inherent to the object. Of course, since we are designing the object, there is nothing to stop us to make our objects to be responding via sonic, visual or tactile cues.

In this article, I address interactivity from the perspective of a new media artist, and discuss its application in the context of interactive art installations. The theoretical framework starts from an embodied music cognition theory paradigm and extends into new media art. As such it answers questions on how interactivity should (or could) be perceived in art and may lead to a new way of addressing the concept in HCI, especially in regards to the uprise of physical computing and novel computer interfaces.

II. How To Read The Map

Art has consistently dealt with mapping problems, although in traditional art they are more related to senses as to sensors. The meaning within an art piece can be regarded as being embedded in the combination of a given sociocultural context, the artist’s concept and the public’s interpretation. How we perceive a piece of art is so to speak embedded in how we interact with our environment and how the individual elements are defined within the piece and the way we interact with them.

How meaning is constructed is a much debated topic in formal and new media art. However in new media art the discussion becomes even more complex when the incorporation of sensors and actuators through the use of technology is included. Including technology enables participation of the public in a valued manner. The degree of participation and the choices for incorporating certain technology become prized parameters in the construction of meaning. In sum it can even be argued that the mapping problem (or more appropriate, how artists handle this) is one of the main topics of new media art. Dealing with the creative or artistic way in which meaning is communicated through form or sound might well be the essence of ‘creating art’.

In interactive art, this mapping problem becomes increasingly complex, where the public is invited to become an equal partner in the construction of meaning. This role is far more than merely fulfilling a role as an...
interpreter of (artistically) implied meaning. Media theorist Andy Cameron addressed the public’s role in interactive media in his presentation ‘Dissimulations, The illusion of interactivity’ [1], he states that “Interactivity is the ability to intervene in a meaningful way within the representation itself, not to read it differently.” [sic].

Lev Manovich subscribes the same idea, in his book ‘The Language of New Media’ [2], when differentiating between ‘open interactivity’ and ‘closed interactivity’. A segregation that is certainly imperative in the context of new media art. However, it is not my intention to give a classification of new media art in which participation of the public is wanted, but to refine the definition of ‘interactive’, making it a more evened out term in research and praxis. Manovich states that the proposition of hyperlinking, one of the key elements of interactive media, “objectifies the process of association” [p. 61] and questions what to make of this “desire to externalize the mind” . There is indeed a danger when following a pre-programmed path, while browsing through a webpage with a fixed branching structure. The choices of the user define the path they follow, and the system they navigate leaves little room to wonder, as it imposes associations upon the user. This is, of course, to a great extent what Cameron refers to as ‘to read it [the representation] differently’[sic]. In contrast, open interactivity can refer to the use of artificial intelligence, artificial life, neural networks and finite state machines, coded into software, leaving room for a more subtle way of interacting with, interpreting and experiencing a work of art. To conclude, he warns about using the term ‘interactive media’ when addressing the post- modern shift towards a physical interaction between the user and a media object, and states that this occurs “…at the expense of psychological interaction.”. I concur with Manovich, although I perceive it as a temporary problem, at least in interactive art. One that can be solved by upholding a fully embodied standard in creating art.

III. Interactivity In Music And Sound Art

Within music research, interactivity is a well established concept, since it used to comfortably reside within psychology. However, in recent years the development of new technology has been prying at this comfortable position. In the past a musical instrument has been seen as a natural extension of the body, enabling the performer to surpass the limitations of the body and give way to new corporeal possibilities. When developing new (digital) instruments, this natural extension is not always a given. While classical instruments are very much modeled around the body, new (digital) instruments are often derived from or build upon existing HCI interfaces. Interfaces that are often more inspired by technological advances in the research field, than an intuitive usability. Wether these instruments are more screen-based or tangible, they usually do require more mental processing to make sense of how the interplay of different media works, both for the performer as the public. This interdisciplinary nature is part of every day life and naturally takes hold of music as well. Although this might complicate the discussion on interactivity in regards to music research, by upholding an embodied approach to (musical) interaction, we pave the way to safeguard the psychological interaction, which is of utmost importance when we are confronted with art.

The embodied view of the process of interactivity is of a cyclic nature. Marc Leman describes it as an actionreaction cycle (fig.1) in his book ‘Embodied Music Cognition and Mediation Technology’ [3].

Interaction is made apparent by using the metaphor of how an instrument is built. While playing the instrument, the resulting sound is processed by the human auditory system. A perception is build up in the mind and judged, by undertaking an action the instrument can be changed. This results in a change of the conditions of the instrument and as a consequence this changes the sound produced by the instrument when it is played.

Leman extends this idea in his model of musical communication, with the purpose of communicating musical intentionality between listener and performer. This is realized through corporeal articulations, transformed through the use of a mediator. The mediation technology should ideally evolve afresh as an extension of the body. This time it should be capable of interpreting the intent of the performer and distinguishing between various sets of actions, by implementing behavior in software. This implies that the mediation technology should be considered no longer merely as an object to which the performer needs to
focus his energy, but as an agent, that is on its own capable of interpreting this energy.

IV. Interactivity In New Media Arts

In new media art, which can include multi sensory aspects and is not only sound-related, the action-reaction cycle or model for musical communication may be less apparent. In his introduction on multimedia environments, Leman proposes to “create an autonomous virtual social agent that is able to communicate...” [sic], such agents should be able to deal with capabilities of both synthesis and analysis. In order to do so, it is necessary to have a mapping strategy that is deduced from objective measurements, ideally cross-referenced with an analysis of subjective experience. This mapping strategy should also have a more universal nature than a mapping available to the public based solely on the artist’s decisions.

V. Affordance

Marcel Duchamp’s exhibiting ‘Fountain’ in 1917 is probably one of the most well established examples to date of the use of an everyday objects outside its known context. When coining the term ‘Ready-Made’, he described it as an object where we have no emotional relationship with, nor a certain opinion about, readymades are an exercise avoiding inurement. Todays technological possibilities make it possible to surpass the level of known things and artists devoutly spend their lives visualizing their fantasies through art. Still they deal with conventions and constraints, whether they are sociocultural or coming from art history, as an important aspect of their creation process and how the public will perceive the work.

Although Marcel Duchamp predates the theory of affordances by a few decades, ‘Fountain’, and any other readymade, is strongly related to it. Affordance is a term the perceptual psychologist James J. Gibson introduced in his 1977 article ‘The Theory of Affordances’ [4]. Gibson meant by affordance ‘an action possibility available in the environment to an individual, independent of the individual’s ability to perceive this possibility’ [5]. Donald Norman introduced the term to the HCI community in his book ‘The Psychology of Everyday Things’, and differentiates from Gibson’s theory, ‘...the term affordance refers to the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used. A chair affords (‘is for’) support and, therefore, affords sitting. A chair can also be carried.’ [6 p.9]. The big difference between both theories is that Gibson addresses affordance from the point of view of perception, while Norman refers primarily to the fundamental properties of an object. A logical distinction if we take their independent goals into account. Gibson was primarily interested in how we perceive the environment, while Norman’s interest lies in manipulating or designing the environment. Taking both approaches into account is of vital importance for the suggested mapping strategy.

A. How Affordance Is Used To Evoke Behavior

In POET (The Psychology of Everyday Things) Norman states that ‘understanding how to operate a novel device has three major dimensions: conceptual models, constraints and affordances’, emphasizing the role of past experience and culture. In regards to design in HCI, Norman makes a distinction between ‘real affordances’ and ‘perceived affordances’, although design is about both, he states that ‘the perceived affordances are what determine usability’ [6, p.123].

Gaver [7] sees affordance as ‘properties of the world that are compatible with and relevant for people’s interactions’ and addresses the common examples as ‘perceptible affordances’. Designing easily-used systems is making affordances perceptible. To Gaver, the concept of affordances ‘implies that the physical attributes of the thing to be acted upon are compatible with those of the actor, that information about those attributes is available in a form compatible with a perceptual system, and (implicitly) that these attributes and the action they make possible are relevant to a culture and a perceiver.’ All of the theories imply action-reaction possibilities, whether invoked by culture, memory or functional design. A well designed object will lead to the user acting upon it, this is certainly true for utilities but is equally true for interactive art installations.

B. How Affordance Is Used In Art

Within interactive art installations great care is taken in designing the installation itself. Installations can be minimal in design or visually exuberant, small or huge, but they all in one way or another expect the user to experience them fully. The designs that gush a myriad of new technological features are often technically esthetic but their function is awkward. On the other hand, designs based on users’ current articulated needs and tasks surpass the potential nested within new technology. This leads to designers introducing terms like ‘intuitive gestures’ when ‘pinching’ photos on their iPod Touch (apple), or ‘throwing a fisbee’ to share selected content and catching the data ‘like catching a ball’ (www.hoccer.com/video.html). Finding new ways of interfacing with novel devices is both exiting and (to some) frightening, and advertising this include a marketing campaign focused on how natural everything appears to be, comparing it to fun games we remember from early childhood. Finding the right balance between new technology and a natural way of interfacing, is therefore imperative, and doing so leads to truly experiencing a work of art or a certain interface. Gaver
gives a similar notion in reference to affordances when stating that ‘Affordances are not passively perceived, but explored...Exploration of afforded actions leads to discovery of the system, rather than knowledge of the system metaphor leading to expectations of its affordances.’

Whether or not we are designing a computer interface, a new music instrument or a fully multi modal artistic environment, we should keep in mind that exploration of a new yet (strangely) familiar world is key for an intimate artistic experience.

VI. Aisthesis And Mediality

Gibson, Norman and Gaver neglect affordances in the context of social interaction. Both Marc Leman and Lev Manovich, as do I, believe that social interaction is of great importance. Especially in new media art, with its remarkable exploratory ways of ‘interfacing’. Hence, the question arises to what extent the actions of others guide our interaction? On a different scale we can ask ourselves if the way we perceive things is not largely influenced by the technology we have been using up until now. According to Jin Hyun Kim et al ‘computer-human interfaces (CHIs) serve as media, not only in the sense of technical apparatus but also in terms of performing inter-medial translations which act as a condition for the emergence of meaning and/or experience’ [8] Creating such interfaces, therefore, requires artistic and technological mediation strategies. She introduces the term ‘mediality’ to deal with the question of how CHIs mediate ‘meaning’ and shape the experience. In Media theory, a medium becomes ‘transparent’ so that the ‘mediatised’ comes to the foreground, as such a medium is characterized as being a sterile empty vessel. When dealing with new media art, however, the interface retains its material presence and stays, for a part, opaque. Even though, often, the artist would like to support the illusion of non-mediation in his work, creating a sense of being there, a feeling of presence.

This seems to be a contradiction, while in fact it is not, if you’d take social interaction into account. The goal of the artist is to communicate intentionality, as introduced by Marc Leman referring to ‘musical intentionality’ or Manovich when stressing the importance of psychological interaction. On the public’s side, the goal is to pick up on the artist’s intentions and distillate a meaning. In return the artist wants to know how his art is perceived by the public, which renders into a cyclic process, key when addressing ‘interaction’

The invitation to act that is embedded in affordance theory might be elaborated within social interaction as an ‘invitation to enact’. This whole process is made possible by different strategies, of which mirroring behavior is the most crucial. The experiments of Meltzoff [9] on imitation behavior in new born human babies have started a silent revolution in thinking about early childhood. Based on these experiments, imitation is believed to be a basic form of corporeal articulation; which is goal-directed and based on purposeful action [10]. Imitation can also be seen as part of a learning process, including a decomposition of the observed sensory action into constituent components, encoded in motor components. Which is followed by a reconstruction of the action pattern from the motor components. This decomposition is guided by an interpretation of the motor pattern as a goal-directed behavior. The main characteristic of corporeal imitation, in regards of imitation of moving sonic forms is, according to Marc Leman, body movement, and is based on a mirroring process which, in turn, is based on both multi-sensory information-processing and the sensing of movement (kinesthesia) [3](Ch.5 p. 110). In the same chapter Leman mentions ‘Embodied Attuning’ [3](Ch.5 p.115), which implies corporeal music in accord with music. Attuning brings the human body into accordance with a particular feature of music, a way of navigating with or inside the music. All of which leads to empathy, the ability to share another person’s feelings or emotions as if they were one’s own (see, e.g., Berthoz and Jorland [11]). This assumes participation, identification and understanding. Recent results (e.g., Carr et al. [12]) suggest that the motor system may access the emotional system with different degrees of engagement, offering a view of how behavioral resonance to affect emotion could be accessed. Embedding emotional intention in a (virtual) agent through the use of a (virtual) motor system is an approach that is a common practice in responsive environments (e.g. Philip Beesley, Hylozoic Soil [13]) and/or interactive art.

In artificial environments, where the artist is only represented by his interactive artwork, the cyclic process, previously mentioned, needs to be implemented within the action reaction possibilities, and even action perception possibilities, of the artwork, which suggest agency. According to Jin Hyun Kim ‘an interface ... is defined as a part of the machine through which it “communicates” with its environment.’ and ‘An interface mediates sensory and motor processes of interacting entities.’ [8]. In interactive installations this would suggest that the interacting entities are both of an organic and inorganic nature. Because of the nature of interaction this would mean that both entities are, to some extent, sentient. Although this term is avoided in AI research (mainly because of the possible ethical dilemmas) the definition of sentient, being ‘able to perceive or feel things’ does attribute to what I believe an interactive art piece should be. Moreover, when we spend so much time and attention to creating anthropomorphic actions, it is a logical step to add the same level on the sensing side. To confirm with AI research, it might be better to imply the installation to
seeming sentient. Doing this implies a mapping strategy that is scalable and dynamic.

VII. Dynamic Mapping Strategy

The development of sensor technology and implementation of this technology in interactive art works, through mapping strategies, should be guided away from the object (or interface) towards the experience. Focus should be on humanizing the objects rather than objectifying humans. Creating interactive artworks, therefore, require a sensitive and variable mapping strategy, where multi-sensory information-processing is valued alongside the installation being aware of its relation to the environment and its own kinesthetic qualities, requiring agency. This should be done with unobtrusive sensor-technology to allow a sense of presence or flow.

The strategy I propose for monitoring behavior can also be used to introduce behavior to artificial entities. It enables us to mimic, or mirror, behavior. This behavior is deducted from objective observations and verified with subjective experiences.

A. How We Can Implement DMS Using FSM

The website of the National Institute for Standards and Technology states: ‘a Finite State Machine (FSM) is a model of computation consisting of a set of states, a start state, an input alphabet, and a transition function that maps input symbols and current states to a next state. The computation begins in the start state with an input string, and changes to new states depending on the transition function.’ [14]. At its simplest, it is a model of behaviors of a system or a complex object, with a limited (finite) number of defined conditions.

![Example of FSM](image)

The states define behavior and may produce actions. State transformations are movements from one statement to another. Such a transition is executed when a certain transition condition is met. Usually this requires an input event, either internally or externally generated, triggering one of the rules that lead to the transition. Entry actions define the initial state, providing a starting point. The FSM is self aware, by keeping track of its current state, remembering the product of its last transition. In figure 2 a simple example of a FSM is given, with only two states, two conditions, transitions and possible entry actions.

The use of FSM originated in mathematics, where they were initially used for language representation. But they were quickly adopted by Artificial Intelligence research because of its apparent simplicity to model behavior. They have been used intensively to model the behavior of foes in first person shooters, such as Quake, which lend its game engine to Unreal Tournament and Half-Life later on.

The original FSM is deterministic; a current state together with a given input would always result in the same, predictable, state transformation. This is not always desired, a next step is creating a non-deterministic FSM. Doing this implies that the state transformation is not (easily) predictable. When multiple inputs are received at various times, each weighing into the transition conditions individually, the outcome is less predictable, making it an event driven system.

Making the FSM even less deterministic, or making a system that upholds its logic but seems to display free will, can be done by introducing Fuzzy Logic. In this fashion a Fuzzy State Machine (FuSM) is created. Here a fuzzy value is assigned to various inputs to represent the degree an input gets defined. The FuSM would take these values into account in regards to state transitions or transition conditions. Conflicts in transition conditions are then taken into consideration by fuzzy logic, determining the outcome transition. The fuzzy logic system uses weighted input values in evaluation of rules, triggering only state transitions above a certain threshold, making the state machine unpredictable. However, this results in a less transparent FSM, which is not always desired. Incorporating random values into a FSM is another approach, but using only random values, makes the use of a FSM more or less redundant.

In regards to interactive art, the bottom-line is that you need to have a right balance between control and exploration possibilities, while interacting with installations. Having control makes the system comprehensible, while having a world of exploration at your disposal makes sure people feel compelled to continue interacting, discovering the installation layer after layer. Having the right balance makes sure people don’t feel lost or intimidated.

B. On What Side

I have mentioned FSM, until now, as primarily targeted towards actions, in that sense the transition conditions are sets of rules and await input of sensors, both internal as external, in order to trigger the state transition. In both human robot interaction, an essential part of interactive kinetic sculptures, as computer human interaction, essential in audiovisual installations that not...
include mechanics, triggering a sensor often makes something move, whether this is a mechanical movement, a moving of pixels on a screen, or pressure waves moving through the space carrying sound. All of which imply a behavior that, within the context of FSM, can be seen as a state change.

It is a common usage for FSM to implement nested behaviors, sub states, and so on, creating high level control with individual outcomes depending on the same input signals. This is especially obvious in single mode of first person shooters, there is only one player but different enemies, all enemies share the same goal (killing the player) and thus react to its presence, but all in different ways, using different behaviors. However, the input conditions are, for a larger part, left out of the equation, rendering the input static.

In interactive art, this should be avoided, in view of the fact that we want our installation to appear sentient. Therefore, a FSM machine should include a variability of reading the sensors, according to the state it is in. The linked behavior of the installation should coincide with a measurable and, to some extent, predictable behavior of the public, making it possible to have a sensible mapping in favor of sentient behavior.

This would require a second FSM to be implemented on the sensing side, taking its transition conditions from the FSM machine on the acting side. The action of this ‘sensing FSM’ would involve filtering, scaling and interpreting the sensor inputs, and parsing them to the ‘acting FSM’. This ‘sensing FSM’ can be addressed in the same manner and with the same precautions as any FSM, making it either fully deterministic or more free.

VIII. Lament: A Simple FSM Implementation

Music centre ‘De Bijloke’ commissioned an installation to accompany the musical program around the Lamentation week organized in 2009.

A. Technology

The installation consists out of five suspended megaphones, that are spread throughout the exhibition in a circular fashion. The five Megaphones are altered to be more suited for the installation, the controls on the megaphones are overridden and the amplification and volume control of both the input (microphones) and output (speakers) is controlled by external software.

The different materials used for the walls and unique room acoustics of the exhibiting space make it necessary to be able to adapt the way the installation reacts to sound input. The software, as seen on figure 4, has independent controls for all five megaphones, and the ability to expand the installation with a sixth megaphone, when installed in larger spaces.

Fig. 3. Installation view of Lament, as exhibited at music centre De Bijloke

Each of the megaphones has two sound layers, one is a continuous ambient sound layer, the other is a singing voice, that is only heard when there is a direct interaction with the megaphone. The volumes of both sound layers, and the sensitivity of the megaphones is initiated on start up and can be adjusted according to the threshold, to distinguish between the two sound layers, is set on the first public viewing as default values, and saved with the software. Additional controls included are for reverberation and delay, to compromise for the distinct acoustics of horn speakers and any artifacts that come from the installation site. The megaphones’ microphones, listen to what goes on in the room, and transfer this information to a computer.

Fig. 4. Control software for Lament, written in max/MSP

B. Implementation as a FSM

Each of the five megaphones is an independent cluster of two FSM, one FSM is listening to inputs from the outside world, adjusting states as to what they are capturing in regards to the state of the acting FSM. The acting FSM, in its turn, listens to inputs from the sensing FSM, and adjusts its behavior accordingly. The states defined for the acting FSM are (a) murmuring and (b) shouting, and on the sensing FSM the states are (a) listening closely and (b) listening afar.

State changes occur in the sensing FSM because it distinguishes between two amplitude ranges, on the one hand there are the subtle differences in surrounding
sounds (listening afar). On the other hand, the second amplitude range is sensed when someone speaks directly into the megaphone, resulting in a state transition to listening closely. On the side of the acting FSM, this leads to a murmuring state corresponding to the listening afar state, or the shouting state which is linked to the listening closely state.

Together, the five atmospheric layers which occur from five simultaneous murmuring states, form a musical backdrop, and a possible maximum of five voices, when all megaphones are in shouting state, combine to a room-filling ever, changing sound installation. Because each of the five megaphones is an individual entity, and taken into account that within each state the external inputs from the microphones is continuously influencing the state behavior, the possible variations are myriad.

IX. Conclusion

The use of FSM for interactive installations makes it possible to introduce Dynamic Mapping Strategies for Interactive art. However, simply implementing this, does not necessarily make the experience more meaningful. The paradigm of embodied music cognition, extended towards new media art, may well provide a way of solving some of the persistent problems in the development of intuitive mappings. Therefore, an extended study with a combined objective approach, measuring sensor data, and a subjective approach, measuring perceived experience, on how people interact with new media art is imperative. Extending this to include monitoring robotic or virtual behavior, using measurements of their senses (sensors) and their experiences (states) is a given, since we model our interactive environments on the knowledge we obtain from HHI. This leads also to the inclusion of affordance theory both from the viewpoint of the subject (Gibson) and the object (Norman), the inclusion of the theory of mediality, since it clearly shows that there is an evolution in thinking and feeling initiated by merely using technology. For artists, this becomes even more apparent, when the concept of their work is guided by the way they are computing.

In discussing Lament, a concrete example is given on how the implementation of a FSM, on both the sensing as the acting side, helps to introduce interactivity. Although the amount of different states available in the Lament installation is limited, it is clear that this evokes an emergence, while keeping the readability of the installation. However, it should be noted that the readability lies not merely in the implementation of the FSM paradigm, but lies within the combination of affordances, mediality and usability.

It is our believe that combining theories from HCI, HRI and HHI, as supported by this paper, will result in installations that envision interactivity in a more social manner, leading to a more valued artistic experience.

Acknowledgement

I would like to thank my promoter Prof. Dr. Marc Leman for his insightful ideas and remarks. My research is made possible by the EmcoMetecca project at IPEM. I would also like to express my gratitude towards Ivan Schepers, for his assistance and guidance in developing the Lament installation. I would like to express my gratitude to Benjamin Buch for the extended discussions on FSM. I would also like to thank Uwe Seifert and Jin Hyun Kim, not only for acknowledging new media as a valid testbed for research, but especially for their time and efforts in talking this through with me. Finally I would like to thank all, together with Lüder Schmidt and Son-Hwa Chang, for making our robot experiments not only interesting but enjoyable as well.

References