Symbolic Objects in a Networked Gestural Sound Interface

Eric Kabisch, * Amanda Williams, Paul Dourish

*Arts, Computation and Engineering (ACE)
University of California, Irvine
Irvine, CA 92697-2775
ekabisch@uci.edu

Donald Bren School of Information &
Computer Sciences
University of California, Irvine
Irvine, CA 92697-2775
amandamw@ics.uci.edu, jpd@ics.uci.edu

ABSTRACT
SignalPlay is a sensor-based interactive sound environment in which familiar objects encourage exploration and discovery of sound interfaces through the process of play. Embedded wireless sensors form a network that detects gestural motion as well as environmental factors such as light and magnetic field. Human interactions with the sensors and with each other cause both immediate and systemic changes in a spatialized soundscape. Our investigation highlights the interplay between expected object-behavior associations and new modes of interaction with everyday objects. Here we present observations on embodied network interaction and suggest opportunities for further investigation in this field.

Author Keywords
Augmented reality; interactive sound; social computing; embodied interaction; sensors; symbolic objects; play.

ACM Classification Keywords
H.5.5 Sound and Music Computing: Methodologies and techniques; H.5.3 Group and Organization Interfaces: Synchronous Interaction.

INTRODUCTION AND MOTIVATION
In ubiquitous computing settings, computation moves "off the desktop," and, by the same token, interaction moves into physical space. Interaction between humans and computers becomes subsumed within, and mediated by, interaction between humans and their surroundings. This transformation is significant, given that conventional user interface design has traditionally exploited metaphors of real-world objects that are, nonetheless, critically distinct from that real world [7].

We use the term "embodied interaction" to refer to forms of interaction with computer systems that are embedded in the physical and social worlds [3]. Artifacts and spaces become meaningful for individuals largely through the way they are used by other social actors; collective social action, both real-time and accreted, configures the ways in which we interact with the physical. So, two critical issues arise for interaction with ubiquitous computing – first, how can people understand and make sense of complex, embodied computational systems, and second, how do these understandings arise out of the collective actions of many participants?

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

We have been investigating these questions through an interactive sound installation called *SignalPlay*, in which participants use physical interface objects to explore a complex auditory environment. SignalPlay takes sensor data collected in real time and uses it to generate and manipulate sounds that are fed back to the participants. The configuration and movement of different objects control aspects of the soundscape; as participants explore the space, they begin to associate their own actions with the response of the system and advance their process of discovery.

Critically, this is a collaborative endeavor; multiple participants share the same space and interact with the system simultaneously. A user's interaction not only affects sounds they can readily associate with their own actions, but it also induces global systemic changes, thereby affecting the sounds generated in response to other participants' actions. This creates a collaborative mesh of interaction where the user is not only engaging in physical action coupled to auditory response, but also where visual cues of co-participants and their behaviors are woven into the fabric of cause-effect associations. SignalPlay is an initial foray into the phenomenology of augmented space.

RELATED WORK
Our investigations were inspired partly by earlier uses of complex audience spaces as a focus for embodied interaction, partly by studies of the collective experience of exhibits and gallery spaces, and partly by considerations of metaphor and instrumentality in interaction.

Art practice has long explored ideas of computational sensory feedback based on physical interaction. These ideas appear in 1950's and 1960's explorations such as a photoelectric and microphone controlled sound system designed by Billy Klüver for a series of performances held in October 1966 under the title *Nine Evenings: Theatre and Engineering* [2]. Installation and performance artists such as Myron Krueger and David Rokeby have continued to explore the use of sensor technologies with real-time sound generation.

There is existing work on the use of gestural user interfaces for electronic instruments including that of Tod Machover and the Hyperinstruments group at MIT Media Lab. The *Beatbug* system [9] in particular focuses on users' ability to manipulate musical system behavior at different levels of collaboration and complexity using simple toy-like objects. In contrast, SignalPlay uses music as a means of exploring a novel interface; we do not think of it as an "instrument," but as an experience. It draws on the idea of tangible bits [6] and is more akin to *ensemble* [1] or the Cardboard Box Garden [4], which use physically embodied audio spaces to investigate the augmentation of familiar objects with computational capabilities; as we will describe, we extend this by exploring how systemic and collective sound controls affect collaborative interaction.

Studies of interaction with exhibits in museums and galleries [5,8] have drawn attention to the ways in which the experience of exhibits is embedded not only in a physical space, but also a space which is inhabited by others, whose actions serve to configure that space for each participant. These detailed studies of interaction informed our observations regarding the use of our system.

DESIGN

SignalPlay was built with Crossbow Mica2 motes running TinyOS, developed at the University of California, Berkeley. They are small enough to embed in our objects, and form ad-hoc networks through radio communication. They are fitted with sensors including accelerometers, magnetometers and light sensors. Each mote regularly transmits sensor data to a receiver mote, attached to a PC laptop. A Java application reads and formats the sensor data and sends it to a Macintosh laptop. Audio content is then generated from the sensor data using object behaviors and music created in Max/MSP and Reason. Spatialized audio is output through a multi-channel sound interface.

Interface Objects

Symbolic objects, and what Smith calls the tension between literalism and magic [7], are central concerns in this design. On the one hand, the objects must, through their physical affordances, suggest how they should be handled; on the other hand, their effect upon a complex audio environment can rarely be conveyed through form alone.

We selected specific objects based on their capacity to elicit prescribed behaviors and their relation to the theme of "play." The objects are three giant chess pieces (a rook and two pawns), five oversized building blocks, two bongo drums, a navigational compass in a wooden box, and a Star Wars lightsaber. Due to the interaction between their physical forms and their programmed behaviors, these objects exhibit different degrees of conflict between literal and magical.

The three chess pieces sit on the ground amid a "chess board" of six disjoint squares, designed to cue the participant to move the pieces around the space. A mote is placed inside each chess piece such that moving and setting down the chess piece triggers its behavior.

Figure 1. Blocks

Five 12" cube building blocks are arranged on and around several small pedestals (see Figure 1). Each has a hole in the top under which a light sensor is placed. The expected behavior of stacking blocks on top of one another drops the light reading below a set threshold and the system responds to that stimulus.

The bongos have holes in the top with light sensors inside each drum. In striking the center of the drums the user affects the light readings and thereby controls a bass line in the system. The system behavior is sensitive to which drum is struck and how long the light source is obscured.

The box-mounted compass is hinged in two directions, allowing it to swivel when tilted. It is fitted with a mote using accelerometer and magnetometer sensors. When the compass is at rest or the compass lid closed it is silent. By opening the lid, the user activates its sound and controls various parameters of a waveform synthesizer by moving, tilting and rotating the compass.

The lightsaber is fitted with a mote mounted to the handle. It sounds upon sensing motion and is silenced after several seconds at rest. When swung by a participant, the speed at which it moves dictates the enacting of sampled sounds.

Direct and Systemic Sound Control

Each object affects the system in a readily apparent way through discrete sound events (*direct controls)* that occur in immediate response to the interaction. In addition, most of the objects have effects on a system-wide level (*systemic controls*), thereby changing the sounds of other objects. In this way, participants begin to engage in a process of interaction not just between themselves and the system, but also indirectly (and directly through social behavior) with other participants.

The systemic control of sound feedback is currently based on control of tonal harmony (keys, scales and intervals), tempo, and timbral quality. For all of the objects except the lightsaber, we base the direct sounds on a globally specified pitch we call the *tonal center*; if the tonal center is changed, these sounds are transposed in pitch by the same interval. These objects, except for the compass, are also governed by a *scale* of specified intervals relative to that tonal center. The object sounds base their tonal harmony on a *set* of pitches defined by the tonal center and scale intervals. However, object sounds are not confined only to pitches in that set, but can also deviate by a chosen interval from specific pitches within the set.

The object sounds form a continuum from simple and direct control to complex and systemic control in the following order: lightsaber, compass, bongos, chess pieces, and blocks. The lightsaber uses only direct controls with no affect on a system-wide level. The compass is affected by the tonal center but not the pitch sets. The rest of the objects have direct controls with an increasing level of system controls. There are also sounds that are not related to the physical objects; these are based entirely on systemic changes and have no direct control.

OBSERVATIONS

We deployed SignalPlay in four showings. The first was an opening event for a new research building at the University of California, Irvine (UCI); the other three were showings at the Alien Art Studio in the Arts, Culture and Technology building at UCI. We video-recorded people's interactions with the exhibit and received informal feedback from them during and after their interactions with SignalPlay. The specific examples discussed in this paper were taken from a session in which naïve participants were allowed to play with the system for 15 minutes. They were then presented with an explanation of the system and allowed to play some more. A full analysis of the collected materials is ongoing. However, a number of broad findings cast light on our fundamental questions concerning symbolic and collaborative aspects of embodied interaction.

Perception of Objects

One critical design concern is the extent to which the physical artifacts encourage natural interaction while also providing a route by which that interaction can be transformed as participants explore the audio space. Certain objects exhibit little tension between symbolic physical cues and augmented behavior; the lightsaber, for example, is enhanced merely to make lightsaber noises when swung like a sword. This object had enormous initial appeal, and was quickly understood.

Figure 2. Compass as musical controller

Other objects support richer modes of interaction. Over the course of five minutes, we observed a man learning to control aspects of the system's sound generation through the compass. Initially, he held the compass stationary in front of his torso and walked around, changing direction periodically. As it became evident that direction influenced the compass's associated sound, he stood still and rotated his torso, holding the compass rigid. This behavior turned to rotating the compass with his hands while holding his torso stationary. He then tilted it, resulting in a dramatic pitch change. With more experimentation, he combined direction with tilt, as well as opening and closing the lid to abruptly start and stop the sound. His interaction with the object departed significantly from ordinary compass use as he learned to understand the augmented compass as a new object in its own right.

Social Interaction

Since the system's controls are distributed amongst multiple objects, group play is crucial to establishing an awareness of system interaction. Systemic controls were most evident when several users were interacting at the same time, triggering objects in quick succession. Certain understandings and behaviors emerged only during heavy collaboration.

The chess set is a case in point: the systemic change in tonal center affected by the rook was most apparent when other sounds were triggered concurrently. Due to the size and dispersal of the chess pieces, one person could not move them rapidly enough to make the effects obvious. Once the workings of the system were explained to the participants, several gravitated towards the chess set, which had previously generated interest only in a couple individuals. These groups remained engaged for longer than the previous solo players and, in attending to the objects' capacity as sound controllers, departed more from the symbolic cues of the chess pieces; illegal moves were made more readily and conventions of turn-taking were discarded (see Figure 3). Throughout the session, participants' understanding and experience of the exhibit was shaped by directly witnessing the actions of others. At the same time, SignalPlay provided its own manifestation of co-participants' actions. In one case, a woman playing with the blocks responded to a sudden change in the compass's sound with an expression of surprise; in another, hearing the bass-heavy *bong* of the rook, a man momentarily turned his attention to two women playing with the chess pieces. Participants' experience of each other was mediated by SignalPlay at the same time as their experience of SignalPlay was mediated by each other.

Figure 3. Speed Chess

CONCLUSIONS AND FUTURE DIRECTIONS

We set out to examine how people can understand physical interfaces to complex systems, in light of Smith's tension between "literal" and "magical" for graphical user interfaces. The critical difference between our physical objects and Smith's graphical objects is that physical objects cannot be separated from their literal component. If graphical interface objects could be charted as points on a continuum between literal and magical, augmented physical objects would have to be represented as lines anchored in their inherent literal form and extending into the magical, leading the eye from one end of the spectrum to the other. Our observations focus on the transition from symbolic to instrumental use of physical objects; a transition that is not inherent in the objects but an emergent property of embodied encounters with them.

An artifact, through its physical form and socially generated meaning, elicits specific actions from people. In a system that employs symbolic interface objects, these implied actions must generate a perceptible response in order to engage people. Our bongo drums, due to issues with system latency, were often unresponsive to participants' initial taps, and failed to engage them. When sufficiently engaged however, participants were inclined to explore deeper interactional possibilities. In instances where gradual or serendipitous departures from the initial interaction generated noticeable system response, our participants readily reformulated their understanding of objects and the modes in which they engage with them.

The social interaction between participants was crucial. Audible feedback from the system made new discoveries evident to anyone within earshot, and people's use of objects was informed by what they had seen others do previously. Practical understandings of the system were formed by the group as a whole.

SignalPlay is an early prototype. We are continuing technical and interactional refinement of the system, as well as our work with the dynamic generation and manipulation of sound as system response. Future enhancements will include the use of network topology as a way to explore coarse-grained spatial interaction. Through our experiments, we hope to examine emergent forms of collective practice in digitally augmented spaces.

ACKNOWLEDGEMENTS

We would like to thank Simon Penny, Jennifer Rode, Bill Tomlinson, and all of our participants.

REFERENCES

1. Anderson, K. 'ensemble': Playing with Sensors and Sound. *Ext. Abstracts CHI 2004*, ACM Press (2004), 1239-1242.
2. Dinkla, S. From Participation to Interaction: Toward the Origins of Interactive Art. *Clicking In: Hot Links to a Digital Culture*, ed. Hershman Leeson, L, (1996) 279 - 290.
3. Dourish, P. *Where the Action Is.* MIT Press. (2001)
4. Ferris, K., Bannon, L. "…a load of ould Boxology!" *Proc. DIS2002,* ACM Press (2002), 41-49.
5. Heath, C., Luff, P., vom Lehn, D., Hindmarsh, J. Crafting participation: designing ecologies, configuring experience. *Visual Communication,* SAGE Publications (2002), 9-33.
6. Ishii, H., Ullmer, B. Tangible Bits: Towards seamless interfaces between people, bits and atoms. *Proc CHI 1997*, ACM Press (1997), 234–241.
7. Smith, R. Experiences With the Alternate Reality Kit: An Example of the Tension Between Literalism and Magic. *Proc. SIGCHI/GI 1987*, ACM Press (1987), 61-67.
8. vom Lehn, D., Heath, C., Hindmarsh, J. Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries. *Symbolic Interaction,* University of California Press (2001), 189-216.
9. Weinberg, G., Aimi, R., Jennings, K. The Beatbug Network – A Rhythmic System for Interdependent Group Collaboration. *Proc. NIME-02* (2002)