"playing_robot": An Interactive Sound Installation in Human-Robot Interaction Design for New Media Art

Benjamin Buch
Institute for Musicology
University of Cologne
Cologne, Germany
bbuch@smail.uni-koeln.de

Pieter Coussement
IPEM
University of Ghent
Ghent, Belgium
pieter.coussement@ugent.be

Lüder Schmidt
Institute for Musicology
University of Cologne
Cologne, Germany
lueder.schmidt@uni-koeln.de

ABSTRACT

In this study artistic human-robot interaction design is introduced as a means for scientific research and artistic investigations. It serves as a methodology for situated cognition integrating empirical methodology and computational modeling, and is exemplified by the installation playing_robot. Its artistic purpose is to aid to create and explore robots as a new medium for art and entertainment. We discuss the use of finite state machines to organize robots' behavioral reactions to sensor data, and give a brief outlook on structured observation as a potential method for data collection.

Keywords

Human-Robot Interaction, Embodiment, Finite State Automata, New Media Art, Dynamic Mapping, Structured Observation

1. INTRODUCTION

playing_robot was presented at the ESCOM 2009 conference as an outcome of a students' work from the International Summer School in Systematic, Comparative and Cognitive Musicology (ISSSCCM-09), which both took place in Jyväskylä/Finland. The installation served as an environment for an observational study in human-robot interaction and as a first step in (artistic) human-robot interaction ((A)HRI) design. (A)HRI design is being developed as an integrated research methodology combining empirical-experimental research strategies with computational modeling in investigations on situated cognition and as an application of embodied cognitive science to human-computer and human-robot interaction in art and entertainment. For the sake of brevity, we use "situated cognition" as an umbrella term for embedded, embodied, and extended approaches to the mind/brain in cognitive science [15]. In such a situated approach robots and concepts from robotics are indispensable scientific tools to study empirically the "musical mind/brain", interactivity and systems interacting with their social and natural environment as well as the logic of the underlying biological mechanisms. Concerning art and entertainment we agree with Bar-Cohen and Hanson's ([2], p. 112) analysis that "Bio-robotics and AI-driven computer animation promise to transform the face of entertainment and art in ways no previous art medium has" and that ". . . artistic explorations effect change in the mindscape and in our cultural information landscape. They [i.e. explorations using robots] can forever change our expectations of what robots are and should be, and what humans are and what we may be in the future." ([2], p. 114; cf. also the contributions in [9]).

To cope with the challenges sketched above we introduce (A)HRI design for the scientific as well as artistic use of robots in New Media Art. The scientific approach uses the metaphor of (A)HRI design to guide experimental design. (A)HRI design is thought of as an extension of observation, experiment, measurement, and modeling to real world situations. Its goal is to establish a methodological foundation for the idea that New Media Art is a testbed for scientific research [17, 18]. The artistic purpose of (A)HRI design is to explore the creation of interactive art and entertainment with robots as a completely new medium. playing_robot is the first installation we developed to exemplify these ideas.

2. *PLAYING_ROBOT* INSTALLATION

2.1 Description of playing_robot

playing_robot is an interactive installation which utilizes embodied agents (robots) equipped with sensory and motor devices and computational sound processing to enable interaction with the visitors. It is the first interactive union of a Lego Mindstorms NXT[1] robot, a Khepera III[2] and humans.

For playing_robot, the Lego Mindstorms NXT construction system was used to build a turtle-like robot ([3], Ch. 6). The Turtle utilizes one control unit based on a 32-bit microprocessor and three servo motors: Two motors actuate the four legs for slow, "crawling" movement; the third one animates the head which consists of a NXT ultrasonic proximity sensor.

The Khepera III robot used in this installation is a small (diameter ca. 13 cm) mobile robot running an embedded Linux operating system on an ARM processor. It is driven by two motors and equipped with ultrasonic and infrared proximity sensors placed horizontally around its body.

The installation takes place in an area of approximately two meters in diameter with no other objects than the robots in it.

When the installation starts, the Khepera III is driving in a circle of approximately 1.5 m in diameter while the Turtle is located a short distance away, facing in the direction of the Khepera III. The Turtle is sitting still, with random leg movements of short duration from time to time. A children's song is played as an artistically inspired choice to contextualize the robots' playful behavior.

[1] http://mindstorms.lego.com/en-us/default.aspx
[2] http://www.k-team.com/

Approaching Khepera III: When a visitor approaches the scene and comes close to the Khepera III, the infrared proximity sensors of the Khepera III are affected. If the distance to the Khepera III falls below approximately 20 cm, both robots change their movements: The Khepera III traces backwards the same circular path, and the Turtle begins to move its legs back and forth rapidly. Those movements produce whizzing sounds due to motor activity. Additionally, the children's song is processed by granular synthesis, resulting in a very noisy sound.

If the visitor keeps a short distance to the Khepera III, following it on its backward course, both robots again change their movement patterns: The Khepera III starts to move back and forth slowly and the Turtle moves forward, slowly approaching the scene. The children's song is again processed by granular synthesis, but this time to produce a purring sound.

If the visitor does not stay close to the Khepera III, the robots fall back into their initial behavior after a certain time. The sound changes back to the original song.

Approaching the Turtle: When a visitor approaches the proximity sensor constituting the Turtle's head to within a range of 20 cm, it starts to move back and forth randomly for 5 seconds. This head movement neither affects the Khepera III nor the sound that is played.

2.2 Technical Setup

A technical goal of the installation was to have the two robots communicate, one notifying the other if it changed its behavior due to interaction; sensor data derived from human interaction with one robot should also be accessible to the other one.

As the Lego Mindstorms NXT's communication is based on bluetooth whereas the Khepera III communicates via wireless LAN, a direct connection could not be established. Instead, both robots were remotely controlled from a set of computers.

The robots' sensor data was sent to the computers for processing which was integrated with sound generation and the production of motor control commands in Max/MSP. OSC was used as a general protocol for the communication among different applications running on the computers.

As OSC is not natively implemented in the Lego NXT software package, we relied on the Java-based LegOSC[3] for communicating with the Lego NXT. LegOSC works as an OSC gateway and allows any software implementing the OSC protocol to exchange data with the Lego NXT using the built-in bluetooth connectivity.

While the Khepera III is capable of wireless LAN communication, there exists no interface to the OSC protocol. A custom interface[4] was used which enables UDP-based access to the low-level read and write commands of the internal Khepera III control using a PureData patch.

Due to performance reasons, robot control and sound generation were distributed among three computers. The first one was running PureDate and communicated with the Khepera III. The second one was running Max/MSP, did the processing of the data and communicated with the Lego Mindstorms NXT via bluetooth. The third computer was running Max/MSP to process the sound.

2.3 Art as Science

A traditional scientific approach to human-computer and human-robot interaction aims at the evaluation of systems with respect to clearly defined problems to be solved and according performance criteria to be met such as accuracy of task completion and time required [14]. In the context of interactive artistic installations, however, such well-defined criteria will be difficult to formulate. Instead, reference is made to general notions of interaction and to the quality of observed / experienced interactions e.g. in terms of emotional affection or the time resp. intensity of involvement (cf. [16]).

To go beyond informal descriptions and to allow for scientific investigation of behavior in interactive artistic contexts, *artistic* human-robot interaction needs to develop a methodology that does justice to the openness of interactions and yet provides for rigorous analysis. As one possible approach we consider the adoption of structured observation [1]. In order to initiate the development of a coding scheme that is appropriate for the observation of behavior in the context of New Media Art informal observations and video recordings were collected during the playing_robot installation.

Within artistic contexts, finite state machines have been recognized as a powerful means to generate complex and interesting interactive behavior [19, 16]. In particular, one concern in the design of playing_robot was to avoid stereotypical system responses in favor of a dynamical mapping of visitors' behavior to robots' reactions and sound changes, which was based on the state of the system. The intention was to keep the visitors within the interaction process.

Moreover, the idea of finite state machine is discussed in the context of modeling (human) behavior within cognitive science [4]. Because of this possible merger between artistic practice and modeling approaches and because of difficulties arising in the interpretation of playing_robot in this context, the main focus in the following discussion will be on notions related to finite state machines and their extensions.

3. FINITE STATE MACHINES

3.1 States and Behavior

The general idea behind finite automata resp. finite state machines is that the behavior of a system, e.g. a machine or an animal, is described as a function of its current (internal) state and its current input.

Notions of state that have inspired the design and implementation of the present installation originate in a variety of contexts and are within these applied at different levels. In classical descriptions of physical systems (cf. [10], p. 3) state is conceived of as the information that is necessary to determine the further temporal evolution of the system under given circumstances. It is captured by variables whose values in the course of time are given by the solution of a set of first order differential equations taking into account environmental conditions. Ideally, these differential equations will be formal expressions of physical laws, so that the description in terms of state variables will not only yield accurate predictions of future states but also provide a kind of explanation of observed sequences of state values [7].

In the formal context of automata theory the explanatory ideal is typically dropped. Descriptions of computing systems refer to state as an internal or m(achine)-configuration [5, 20] which can change according to certain rules given a specific input and together with the input determines the machine's further operation. Thereby (abstract) sequences of states are separated from their physical realizations and time as parameter is replaced by serial order.

Reference to internal configurations is omitted in an approach advanced e.g. by Minsky [12]: Taking state to be determined by previously encountered conditions and determining future operation of the machine together with

[3]by J. Cardoso, http://diablu.jorgecardoso.eu/
[4]written by T. Grewenig and R. Becker

external conditions, a definition of state is offered as equivalence class of sequences of previous conditions ("histories") giving rise to identical sequences of future operations for all possible future sequences of external conditions.

A related perspective is taken e.g. in the field of behavior-based robotics (cf. Murphy [13], Chapter 5.1). Observable sequences of operations (behaviors) of the systems to be designed are assumed to stand in a one-one relation with internal configurations of the systems, and therefore are taken as unique indicators of system state. In consequence, state and behavior are not treated separately. Instead, states are labeled by the associated behaviors.

In all of these cases, state is employed within a description of a system operating in a certain environment, and it is assumed that a distinction can be drawn between system and environment. Minimally, reference will be made to the possible states of the system, the relevant environmental conditions interpreted as inputs to the system, and system operation via the combined effects of state and input on future states of the system.

The uses of state and related system descriptions range from detailed accounts of physical processes incorporating explanatory aspirations to heuristic tools for the structured description of a situation in the process of system design. These positions are connected by the common reference to formal properties which are investigated in automata theory. Whereas scientifically oriented applications will probably exhibit an inclination towards the former position, in the case of artistic installation design emphasis will be on the aspect of structured description. It is one of the ultimate goals of the approach presented here to integrate scientific and artistic ideas in a coherent and rigorous manner.

The tasks for interaction design faced in the context of the installation presented included decomposition of the situation in terms of the formal elements of the system description (inputs, states, behaviors, outputs; identification of system and environment) and development of satisfactory input – state/output/behavior mappings.

More theoretical considerations will address the adequate choice of formalism and the applicability of formal automata/ system theoretic results as well as the interpretation of the chosen formalization with respect to the relevant scientific background.

3.2 FSM: Formal Definitions

Generally, two variants of finite state machines are discussed: FSMs which do not produce any output, called *acceptors*, and FSMs with output called *transducers*. If the new state of a FSM is uniquely determined by the current state and input, it is called *deterministic*. In the following, we will restrict ourselves to deterministic transducers.

Formally, a finite state *transducer* can be described as a 6-tuple (cf. [8], pp. 42, 43) $M = (Q, \Sigma, O, \delta, \lambda, q_0)$ consisting of a finite set of states Q, a finite set of inputs Σ, and a state transition function $\delta : Q \times \Sigma \rightarrow Q$.
The output function λ can be defined in two variants:
1) $\lambda : Q \rightarrow O$ associating output only with the current state of the FSM; in this case the FSM is called a Moore machine.
2) $\lambda : Q \times \Sigma \rightarrow O$ associating output with both a state and an input encountered while the machine is in this state; a FSM conforming to this definition is called a Mealy machine.
Finally, an initial state q_0 needs to be specified.

3.3 playing_robot as Finite State Machine

In the case of playing_robot state was used to refer to combinations of concurrently running processes (cf. Section 2.2) which were taken to realize the behavioral patterns described above. Moreover, these states were labeled as playful ($\mathcal{P}$), anxious ($\mathcal{A}$) and trustful ($\mathcal{T}$) in accordance with the artistic scenario alluded to in Section 2.1.

The only inputs to the system *as a whole* were the readings of the infrared sensors of the Khepera III robot. These were differentiated by the conditions of at least one value exceeding 300 indicating an object in the vicinity, referred to as near (n), or all values being less than 300, in the following referred to as far (f). As the conditions for changing the behavior/state of the system also involved the time span during which a sensor reading n did or did not occur, a timing device needed to be included, which was (re-)started every time the condition n was encountered and expired after a period of 15 seconds.

In terms of finite state machines, the situation may be described as follows: The set Q of states contains the three elements represented by the labels $\mathcal{P}$, $\mathcal{A}$, and $\mathcal{T}$, i.e. $Q = \{\mathcal{P}, \mathcal{A}, \mathcal{T}\}$; the initial state q_0 of the system is $\mathcal{P}$.

The set of inputs Σ needed to achieve state changes will contain the conditions n and f as well as a condition indicating that the timer has expired, referred to as timeout (t); in consequence, the timer will have to be considered as external to the finite state machine. In symbolic form we have: $\Sigma = \{n, f, t\}$.

Starting resp. re-starting the timer can be captured by allowing the finite state machine upon (re-)entering the states $\mathcal{A}$ and $\mathcal{T}$ to produce an output start_timer, represented by the symbol $\odot$. Thus, a set of outputs containing this one element will be included: $O = \{\odot\}$.

Because the output according to this scheme is associated with state transitions and thus with combinations of states and inputs, the finite state machine realized here will best be understood as resembling a Mealy-type FSM.

Although it provided inspiration and guidance in setting up playing_robot, the formalism of finite state machines may not be the optimal choice for the purpose of installation/interaction design: As illustrated by the description of the timer above and the missing integration of the Turtle's head movements into this scheme, the formalism does not lend itself naturally to the treatment of concurrently running processes influencing each other within a single coherent system description. Moreover, the processes underlying the different behaviors of the system are indiscriminately lumped together under one label, although in the process of implementation they were treated rather independently. Possibly a more satisfactory framework can be provided by an extension of the FSM approach such as that introduced by Harel (e.g. [6]) under the name of statecharts. These allow e.g. for hierarchical grouping of states as well as simultaneously being in different ("orthogonal") states, and offer the explicit inclusion of temporal conditions.

4. CONCLUSION

Concerning the artistic goal to structure and organize interactive behavior of installations, finite automata appeared to be a good starting point. The FSM we used, however, lacked the features of hierarchical structure and concurrency, which are deemed to be important to achieve interactive or reactive behavior in New Media Art. Therefore, as a next step the exploration of Harel's statecharts [6] – designed to incorporate these features – may be promising. Moreover, from the development of the installation playing_robot arose the necessity to clarify the interpretation of the concept "state" and its relation to the observable behavior of a system.

In the playing_robot installation, part of the visitors apparently became involved beyond the point of merely understanding system behavior, exhibiting what may tentatively be called playful interaction. The structure of installation behavior appears to be reflected in visitors' behavioral patterns, although for a more rigorous comparison the tools still need to be adapted. In developing a coding scheme for structured observation which may serve this purpose, hints may be drawn from the observation that some aspects of behavioral patterns implemented in the robots were also displayed by the visitors, such as backing away from each other and avoiding close contact or gently approaching one another. Moreover – and in contrast to previous attempts at observational studies of human-robot interaction, e.g. [11] – it should be taken into account, that visitors' behavior was not only directed at the robots, but also included co-visitors. Ultimately, an interpretation of these behavioral patterns in a broader ethological context appears desirable.

Three observations of general interest concern the initialization of the human-robot interaction process, the functional role of the robots' appearance, and the importance of sound for contextualizing the situation. The human-robot interaction should be initiated by the robots or more generally by the designed system. This might be achieved by the physical appearance in connection with meaningful movements. An appearance or behavior vestigially being animated may serve to elicit empathy in humans. Animorphic structure of a robot in connection with some "provoking" (sound) activity directed towards the visitors of an installation seems useful to attract their attention, interest, and to stimulate interaction. In playing_robot the robots' identity and the context of the installation were adapted to the robots' physique and not optimal for human-robot interaction: the robots should not be too tiny and if necessary this should be compensated by artistic means. Sounds may serve as a tool to contextualize meaningfully a situation or the state of the system in order to facilitate "communication" with humans.

In the near future, (A)HRI design in connection with neuro-, social and evolutionary robotics will become indispensable for testing and developing concepts in research on the social human mind/brain and its underlying mechanisms as well as for interactive artistic and entertainment applications.

5. ACKNOWLEDGMENTS

We would like to thank Erika Donald (McGill University, Montreal) and Jan Richter (Universität Hamburg) as co-workers, Mikko Leimu (University of Jyväskylä) for technical support, and the organizers of ESCOM 2009 for giving us the opportunity to present the installation.

6. ADDITIONAL AUTHORS

Uwe Seifert, Son-Hwa Chang, Jochen Arne Otto (Institute for Musicology, University of Cologne, Germany)

Jin Hyun Kim (Cluster of Excellence "Languages of Emotion", Freie Universität Berlin, Germany)

7. REFERENCES

[1] R. Bakeman and J. M. Gottman. *Observing Interaction. An Introduction to Sequential Analysis, 2nd Edition*. Cambridge University Press, Cambridge, 1997.

[2] Y. Bar-Cohen and D. Hanson. *The Coming Robot Revolution: Expectations and Fears About Emerging Intelligent, Humanlike Machines*. Springer, New York, 2009.

[3] D. Benedettelli. *Creating Cool MINDSTORMS NXT Robots*. Apress, Berkeley, 2008.

[4] A. W. Burks. Logic, Computers, and Men. *Proceedings and Addresses of the American Philosophical Association*, 46:39–57, 1972-1973.

[5] N. Chomsky. Formal Properties of Grammars. In R. D. Luce, R. R. Bush, and E. Galanter, editors, *Handbook of Mathematical Psychology, Vol. II*, pages 323–418, New York, 1963. Wiley.

[6] D. Harel. Statecharts: A Visual Formalism for Complex Systems. *Science of Computer Programming*, 8:221–274, 1987.

[7] C. G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. The Free Press, New York, 1965.

[8] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation.* Addison Wesley, Reading, MA, 1979.

[9] C. Lischka and A. Sick (editors). *Machines as Agency. Artistic Perspectives.* University of the Arts Bremen Press Series 04. transcript Verlag, Bielefeld, 2007.

[10] A. Messiah. *Quantum Mechnics.* Wiley, New York, 1958.

[11] M. P. Michalowski, S. Sabanovic, and H. Kozima. A Dancing Robot for Rhythmic Social Interaction. In *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction HRI07*, pages 89–96, 2007.

[12] M. L. Minsky. *Computation: Finite and Infinite Machines.* Prentice-Hall, Englewood Cliffs, NJ, 1967.

[13] R. R. Murphy. *Introduction to AI Robotics.* MIT Press, Cambridge, MA, 2000.

[14] K. L. Norman. *Cyberpsychology. An Introduction to Human-Computer Interaction.* Cambridge University Press, New York, 2008.

[15] P. Robbins and M. Aydede. A Short Primer on Situated Cognition. In P. Robbins and M. Aydede, editors, *The Cambridge Handbook of Situated Cognition*, chapter 1, pages 3–10. Cambridge University Press, Cambridge, 2009.

[16] D. Rye, M. Velonaki, S. Williams, and S. Scheding. *Fish-Bird:* Human-Robot Interaction in a Contemporary Arts Setting. In *Proceedings of the 2005 Australasian Conference on Robotics and Automation*, Sydney, Australia, December 2005.

[17] U. Seifert. The Co-evolution of Humans and Machines: A Paradox of Interactivity. In U. Seifert, J. H. Kim, and A. Moore, editors, *Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations*, pages 9–23. transcript Verlag, Bielefeld, 2008.

[18] U. Seifert and J. H. Kim. Towards a Conceptual Framework and an Empirical Methodology in Research on Artistic Human-Computer and Human-Robot Interaction. In I. Pavlidis, editor, *Advances in Human-Computer Interaction*, pages 177–194. In-Tech Education and Publishing, Vienna, 2008.

[19] G. Trogemann and J. Viehoff. *CodeArt. Eine elementare Einführung in die Programmierung als künstlerische Praktik.* Springer, Wien, 2005.

[20] A. Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, Series 2*, 42(3-4):230–265, 1936.