View metadata, citation and similar papers at core.ac.uk

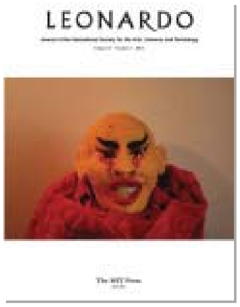
brought to you by CORE

provided by espace@Curtin

PROJECT MUSE®

Evoking Agency: Attention Model and Behavior Control in a Robotic Art Installation

Christian Kroos, Damith C. Herath, Stelarc



Leonardo, Volume 45, Number 5, 2012, pp. 401-407 (Article)

Published by The MIT Press

For additional information about this article

http://muse.jhu.edu/journals/len/summary/v045/45.5.kroos.html

Access provided by Curtin University of Technology (13 May 2014 21:15 GMT)

ARTISTS' ARTICLE

Evoking Agency: Attention Model and Behavior Control in a Robotic Art Installation

Christian Kroos, Damith C. Herath and Stelarc

ABSTRACT

Robotic embodiments of artificial agents seem to reinstate a body-mind dualism as consequence of their technical implementation, but could this supposition be a misconception? The authors present their artistic, scientific and engineering work on a robotic installation, the *Articulated Head*, and its perception-action control system, the Thinking Head Attention Model and Behavioral System (THAMBS). The authors propose that agency emerges from the interplay of the robot's behavior and the environment and that, in the system's interaction with humans, it is to the same degree attributed to the robot as it is grounded in the robot's actions: Agency cannot be instilled; it needs to be evoked.

BACKGROUND AND MOTIVATION

"All too soon we are seduced by Descartes' vision: a vision of a mind as a realm quite distinct from body and world. A realm whose essence owes nothing to the accidents of body and surroundings. The (in)famous 'Ghost in the Machine,'" writes Andy Clark in the preface to *Being There* [1]. At first glance, speaking of agency in the context of robotics seems to pay tribute to this separation of "mind" on the one side and body and world on the other side. It seems to conjure the "Ghost in the Machine" once again. The agent—the one who is driving, leading, acting—is a distinct entity that has been put into the machine. By the same token, it could also be removed and installed in a different machine. In fact, this has been the prevalent view in the past with regard to both forms of embodiment used in the work reported here: Robotic agents and so-called Embodied Conversational Agents ("talking heads"), with the embodiment of the latter restricted to virtual reality but the interaction with humans extending to the physical world. In terms of the technical realization of such agents, the separation suggests itself: There is an input side (sensing) comprising dedicated routines, there is an output side (movements, real or virtual) containing its own control system, and there is something in between that does the "thinking." Conceptually this simplifies the research and implementation work enormously.

This initial impression might be misleading, however. The fact that the agent has to be implemented with modular subsystems handling input and output does not necessarily imply that its inner workings detach the agent from its environment. Similarly, the fact that the agent itself is realized as a modular entity does not necessarily mean that it is driven by the abstract reasoning systems that Clark criticizes. Perception-action systems are able to overcome the modularity suggested by technical requirements through the way they themselves are interconnected. Even if such a system does have a central control system, it need not be a decoupled entity. The degree of its interconnectivity depends on how closely it interacts with other subsystems. For instance, based on the input from the sensors, an attention subsystem might change the properties of a central control system, which in turn might result in different task priorities being forwarded to the attention subsystem. The question of who is driving the agent, the agent within the agent, is exposed as an unhelpful recursive affair. This is also the reason why we speak of "evoking" agency. The agent is not considered something that is in the machine, like a homunculus, controlling it; agency emerges from the interplay of the environment, including other agents, and the machine. In the interaction with humans, agency is grounded in the agenda of the agent to the same degree as it is in the attribution of agency by the human. Most of all, however, we argue, it is grounded in the dynamics of the interaction itself.

We have previously proposed [2,3] that meaningful interactions and the perception of the machine as an intentional agent will occur only if the machine's perceptual and action systems are tightly coupled in much the same way as perception and action are closely linked in humans, according to several psychological theories [4,5]. Due to its important role in a tightly coupled perception-action control system, an attention model has become a central element in our interactive robot, the *Articulated Head*—an art, science and engineering collaboration (Color Plate D). On the artistic side, the *Articulated Head* is based on previous artwork, the *Prosthetic Head*; on the scientific and engineering side, it is based on research and development in the *Thinking Head* project [6]. The *Prosthetic Head* is an automated, animated and reasonably informed artificial head that speaks to the person who interrogates it. Conceptually, it can be categorized as an Embodied Conversational Agent (ECA). However, unlike most of its virtual colleagues, it does not have a specific role to fulfill (e.g. providing information about the exhibits in a museum) but engages in conversations that are in principle entirely unconstrained. The *Prosthetic Head* is not an illustration of a disembodied intelligence. Rather, it raises questions of awareness, identity, agency and embodiment.

There were several reasons for developing the *Prosthetic Head* further into a robotic installation:

Christian Kroos (scientist), MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, Penrith South DC NSW 1797, Australia. E-mail: <c.kroos@uws.edu.au>.

Damith C. Herath (engineer), MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, Penrith South DC NSW 1797, Australia. E-mail: <d.herath@uws.edu.au>.

Stelarc (artist), MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, Penrith South DC NSW 1797, Australia. E-mail: <stelarc@stelarc.org>.

See <www.mitpressjournals.org/toc/leon/45/5> for supplemental files associated with this issue.

© 2012 ISAST

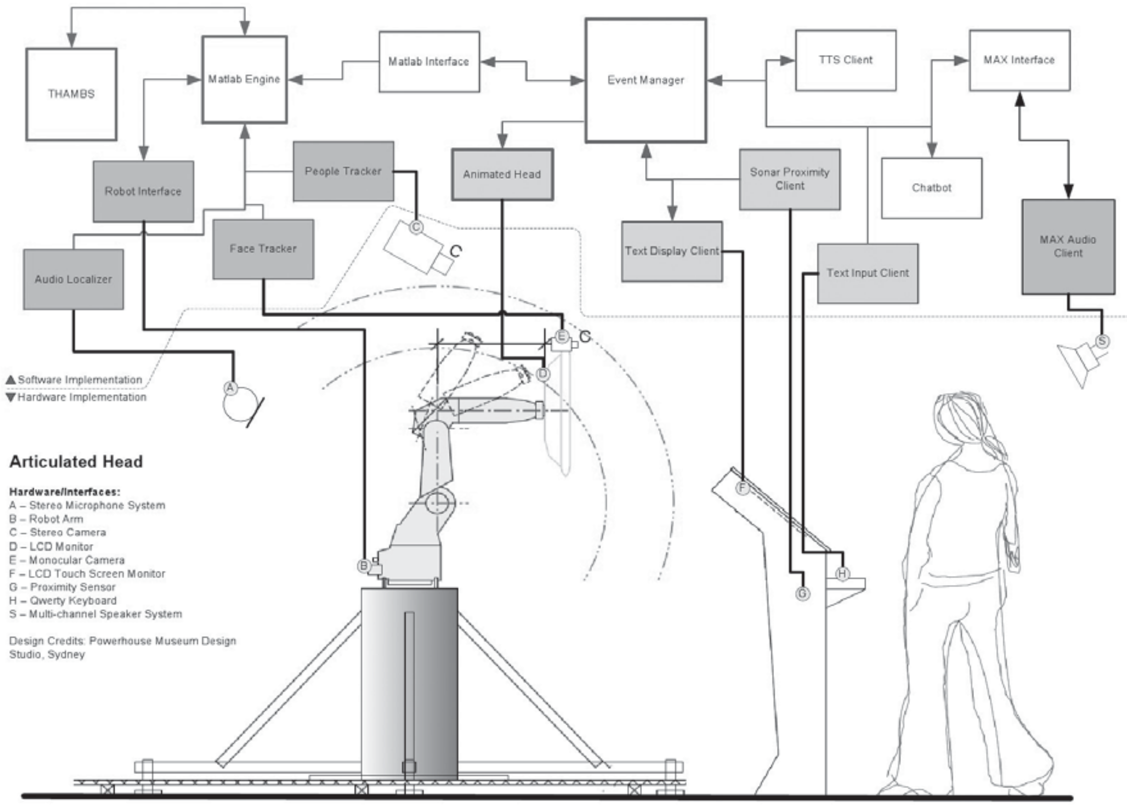


Fig. 1. Schematic of the hardware and software system of the *Articulated Head*. (© Damith C. Herath. Design: Powerhouse Museum Design Studio, Sydney.)

(a) The *Prosthetic Head* was a 5-m-high projection on a wall. Although impressive in scale, it was essentially a screen-based installation and thus purely virtual. In contrast, the *Articulated Head* is an actual-virtual system with an LCD screen imaging the Head mounted to the end of an industrial robot arm that becomes an articulated 6-degrees-of-freedom neck. The fusion between physical and virtual elements is reinforced by the syncing of the physical behavior of the robot and the facial behavior of the *Prosthetic Head*. The robot arm, mounted on a steel base, gives it an anthropomorphic scale and feel. The system is minimal and clean in its aesthetics.

(b) The robot system has a physical, sculptural presence that allowed us to actualize and evaluate sound location, vision tracking and a face-tracking system in a 3D space.

(c) The hybrid robot system bypasses any "uncanny valley" [7] issues, as it does not resemble a human but clearly announces its machine character.

(d) Use of an industrial robot arm ensures that the system performs reliably and robustly.

On the scientific and engineering side, the concept of the *Articulated Head* posed unique challenges. The context of a work of art meant that (a) there would be no clearly defined task, (b) there would be few boundary conditions constraining the interaction with the visitor and (c) expectations with which visitors would approach the *Articulated Head* would vary widely.

Most importantly, however, the *Articulated Head* would have to be perceived as an intentional agent based on its motor behavior alone, in order not to undermine its conversational skill realized via the A.L.I.C.E. chatbot [8] integrated into the *Prosthetic Head*. Perceived agency might not be difficult to evoke, as humans ascribe agency quickly, but the illusion breaks down quickly, too, and evaluation might be unwieldy. If the behavior of the robot appears to be the mechanical consequence of whatever the human user does (or a particular aspect of it), e.g. pursuit based on simple motion tracking, it will be exposed as such quickly; if the behavior does not appear to be connected with the actions of the user, or only insufficiently so, the system will be considered faulty or viewed as random (Waytz et al. [9] discuss some of the properties of human-robot interactions that influence the attribution of intentionality).

As described above, we assume that the solution is to be found in an action-perception control system with a tight coupling between action and perception and an attention model at its core.

Human attention is typically investigated in controlled psychological experiments focusing on specific aspects of the overall phenomenon, e.g. shifts in visual attention triggered by priming stimuli. In thousands of studies, many insights have been gained, yet a general definition of attention has remained elusive. Rather broad primary characteristics have been found to be selection (of sensory information), binding and limited capacity [10]. For attention systems in machines, however, an important distinction between two different types of attention emerged: saliency in the perceptual input (bottom-up or exogenous attention) and task-dependent attention direction (top-down or endogenous attention) [11]. Bottom-up attention can be modeled based on human gaze data obtained with eye-tracking technology. Top-down attention, however, involves high-level world knowledge and understanding and thus largely eludes computer-based modeling. To make things worse, top-down mechanisms appear to be critical, as can be seen in the fact that, even for a barn owl, only 20% of attentional gaze control could be explained by low-level visual saliency [12].

The huge interest in human attention finds a rather small complement in modeling attention in artificial agents. In the majority of cases, attention models were investigated in virtual environments [e.g. 13–16], avoiding problems of real-world object recognition and noisy real-world sensing. A few attempts have been made to develop attention models for robots [e.g. 17–20]. The best known is probably the visual attention system of Breazeal and Scassellati [21] used with the robot Kismet. The attention model presented in this paper differs from these models in that it is more abstract. Low-level salience is provided by the tracking and localization routines and only re-evaluated in the context of our attention model.

There is another important point to be made: In our work we aimed from the beginning to have the robot's behavior emerge from the interaction of its control system with the environment. We avoided pre-scripted behavior as much as possible. Instead of implementing a state-based system governed by *if-then* rules, we opted for a set of subsystems influencing each other through a range of variables and parameters that are dynamically changed by sensory input. What seems to be a minor difference in implementation leads in a few steps from a more or less context-insensitive stimulus-response system to a complex, dynamic system. As a consequence, the *Articulated Head*'s behavior becomes increasingly difficult to predict, something that might be less favorable in most application contexts but definitely not in the case of an interactive artistic installation.

THE *ARTICULATED HEAD*

The *Articulated Head* consists of an industrial robot arm with an LCD screen as its end effector, i.e. the monitor is mounted on the robot arm where in industrial production a tool would be attached. Multiple sensors, including stereo vision, monocular vision, audio and sonar sensors, are mounted on the enclosure as well as on the robot. These sensors provide the necessary "situational awareness" for the robotic agent. An event-driven software framework provides the communication channel between the robot, the sensors and its behavioral control system, the Thinking Head Attention Model and Behavioral System (THAMBS). See Fig. 1 for a schematic of the entire system, which will be described in the following. Note that technical details are omitted here and can be found elsewhere [22].

The Robot

The robot arm, a Fanuc LR Mate 200iC, is a small-scale, highly dexterous and fast-moving industrial platform that has six degrees of freedom (see lower-left side of Fig. 1). It is mounted on a custom-made four-legged structure to provide stable operation. In an earlier version, the robot was enclosed in an octagonal transparent polycarbonate frame. For its current exhibition [23], a new, triangular enclosure was built, consisting of a wooden support and an uninterrupted glass front along the two sides of the triangle (the last side contains a lockable glass door and a small laboratory area for evaluation purposes behind a wooden back wall). In both cases the arrangement serves to maintain good visibility for the observer while preventing users from inadvertently moving into the robot's work envelope.

In order to achieve real-time interactivity and fluidity of motion, the standard interface of the robot has been modified

Fig. 2. Text interface and THAMBS real-time display at the SEAM 2010 exhibition. (© Christian Kroos, Damith C. Herath and Stelarc)

to accommodate real-time motion data that are fed through THAMBS. The robot arm is designed for factory automation tasks in which movements are pre-programmed prior to the production run, whereas in the *Articulated Head* no pre-planned movements or locations are employed. A LAN (Local Area Network)–based interface was developed for this purpose, with additional electronics and interlocks for maintaining safety of operation for both humans and the robot (see "Robot Interface" in the upper-left corner in Fig. 1).

Sensing

Two commercially available camera systems were installed for tracking people in 3D and faces in close proximity (see center of Fig. 1). First, a stereo camera mounted rigidly on the enclosure or on the opposite wall looks downward into the interaction space of robot and visitors. Tracking software ("People Tracker" in Fig. 1) returns localization and height information of all people within the camera's field of view, with considerable tolerance of occlusions and occasional disappearance of the tracked person from the camera's view.

Second, a monocular camera mounted above the top edge of the LCD screen provides the robot with a first-person dynamic view of its environment. Since humans interacting with the robot are of utmost importance for the system, data from this camera are sent to a commercial face-tracking algorithm ("Face Tracker" in Fig. 1). The software routine is able to detect and track a single face in the camera's field of view and returns the face's location and orientation coordinates in the (relative) image coordinate system.

On the acoustic side, the instantaneous location (azimuth) of a moving interlocutor is made available to the system using stereo microphones mounted to the robot enclosure and an acoustic localizer software routine ("Audio Localiser" in the left-middle part of Fig. 1).

In addition to the above components, various ancillary components support the diverse interactive aspects of the *Articulated Head*. A keyboard input device integrated into an information kiosk with an embedded monitor enables text-based interaction with the *Articulated Head* (Fig. 2), and a proximity detector alerts the system to the presence of visitors close to the information kiosk (both in the lower-right part of Fig. 1). A text-to-audiovisual-speech system provides the virtual talking head with realistic speech acoustics and facial motion, and a dialogue management system handles the flow of text input and speech output.

The Thinking Head Attention Model and Behavioral System (THAMBS)

In the *Articulated Head*, the Thinking Head Attention and Behavioral System (THAMBS) manages all interactions and generates appropriate responses. THAMBS consists of four modular subsystems: (1) a perception system, (2) an attention system, (3) a central control system and (4) a motor system. THAMBS is depicted in the upper-left corner of Fig. 1 relative to the entire system, while its inner workings are shown as a diagram in Fig. 3 and are described in the following sections according to the layout in the diagram.

Perceptual Processing

The input received from the sensing routines varies substantially in its form and content, e.g. the acoustic localization software returns an azimuth and a confidence value, while the people-tracking software returns an identity marker and the full set of Cartesian coordinates for each person. To handle this variability, the interfacing routines are set up as "senses" within THAMBS; each comes with a set of parameters and rules controlling the interpretation of the received data values. The perception system of THAMBS transforms the input event into a standardized "perceptual event." It thereby filters out events that do not meet the eligibility criteria set for each sense individually—e.g. acoustic location events with a confidence value below a certain threshold will be discarded. The perception system also receives input about the current state of the robot (angle values of its joints, working status) similar to proprioception in humans and animals.

Attention Model

Very much at the heart of the THAMBS perception-action control system is a biologically inspired attention model. The attentional processing begins with an attention-specific thresholding on the data values of the incoming perceptual events, that is, events with values that do not lie within a pre-determined range will be excluded from further processing. These thresholds are modified dynamically in THAMBS. For instance, when THAMBS switches into sleep mode due to lack of environmental stimuli, the thresholds for the acoustic localization are increased (while all visual input is completely switched off), thus making it more difficult for an acoustic event to reach any further processing stages and "wake up" the system. A perceptual event that passes the threshold test generates as a first step an attention focus.

An attention focus is characterized by its attentional weight, a decay function,

Fig. 3. Diagram of THAMBS, the Thinking Head Attention Model and Behavioral System. (© Christian Kroos)

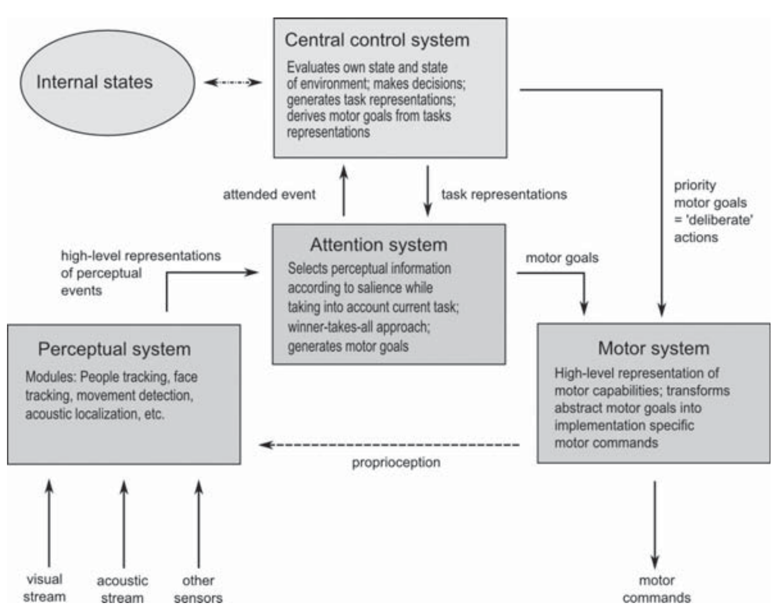


Fig. 4. Within the enclosure of the *Articulated Head* at the Powerhouse Museum, Sydney. (© Christian Kroos, Damith C. Herath and Stelarc)

its lifetime and the spatial location in the real world it is referring to. Its weight is originally determined using a base weight assigned to the type of perceptual event that is the source of the attention focus. Thus, for instance, a face-detection event will receive a higher base weight than an acoustic localization event, as the *Articulated Head* is set up to be geared toward face-to-face interactions with humans. A factor dependent on a chosen property of the perceptual event, e.g. a confidence value returned by a sensing routine, is then multiplied by the base weight to determine the final attentional weight. This is used to increase the stability of certain behaviors, e.g. when the *Articulated Head* is engaged in a face-to-face interaction. If the face-tracking routine returns a high confidence value, signaling detection of a face with the monovision camera, the *Articulated Head* becomes difficult to distract from this interaction and may start to mimic the user's head orientation.

The decay function ensures that an attention focus has a certain lifetime after the event that caused it has disappeared, but that at the same time its strength fades even if registration of the perceptual event is sustained. We used a specific exponential function called the Kohlrausch function, which is known to be able to describe a wide range of physical and biological phenomena [24]. The free parameters of the function are initialized dependent on the type of perceptual event, but, again, they are modified dynamically during run time. Adjusting a stretching/compressing parameter of the decay function toward shorter decays can cause the *Articulated Head* to appear very nervous, constantly switching to new attention foci, whereas adjusting it toward longer decay times will make it appear slow and unresponsive.

The attention foci are in general spatially organized—that is, they are defined via a segment of 3D space centered on the location of the event that attracted attention (compare space- versus object-based attention in models of human attention [25]). This becomes particularly important when the attention system has to determine whether a new perceptual event encountered is—per definition—identical to one of the already existing attention foci. Locations of existing attention foci are matched with the locations of new candidates. If an incoming event and one of the attention foci are indeed found to be identical, the old focus is maintained but its location and weight are updated. The combination of the new and old weights is modeled supra-additively—that is, the resulting value is smaller than the sum of the two original values. The decay function, however, will not be reset in the fusion of attention foci. Thus, even if new events are constantly reinforcing an old attention focus—for instance, a person standing still within the visual field of the *Articulated Head*—the focus will eventually reach very low weight values and will be removed (modeling adaptation). More generally, the above settings enable the *Articulated Head* to strike a balance between focusing on a single source and distributing attention over several sources. In particular, if there is a crowd of onlookers in front of it, it will switch between two behaviors: fixating on a particular person for a while and scanning other people from time to time.

Once all attention foci are created and the decay of their weights computed, one of them is selected as the sole attended event. This is usually accomplished using a winner-takes-all-strategy—the attention focus with the highest weight is chosen.

Finally, the attended event is sent to the central control system. In addition, the attention system creates a motor goal *look_there*. It is designed to point the LCD monitor toward the spatial location of the attended event in order to create the impression that the virtual head shown on the monitor is looking at the event that grabbed its attention. This entails that the monovision camera mounted on top of the monitor is also directed toward the attended event.

Central Control System

The primary role of the central control system is to generate a response behavior appropriate to the constantly arriving input, which, in turn, is affected by this very behavior. The response generation is realized as a non-trivial stimulus-response system—non-trivial because the conditional rules governing it are modified during execution time and are at some points subject to probabilistic evaluation. The conditional rules are called *behavior triggers*. Most behavior triggers result, if activated, in a motor goal that is passed on to the motor system. However, other behavior triggers only change internal variables (such as the attention base weights) and modify the impact of future sensory information or the way certain motor goals are executed.

Motor Control

Motor goals are abstract representations of motor actions to be executed by the robot arm or the virtual avatar displayed on the monitor. The motor system is responsible for converting the abstract motor goals transmitted from both the attention system and the central control system into concrete motor commands or primitives. At first, the motor system determines which one of the two motor goals—if both are in fact passed on—will be realized. In almost all cases, the "deliberate" action of the central control system takes precedence over the pursuit goal from the attention system. Only in the case of an event that attracts exceptionally strong attention is the priority reversed. In humans, this could be compared with involuntary head and eye movements toward the source of a startling noise or toward substantial movement registered in peripheral vision.

The motor subroutines request sensory information if required for the realization of the motor goal, such as the location of a person to be "looked at." They then transduce the motor goal into motor primitives—that is, in the case of the robot arm, into target angle specifications for the six joints.

Performance in Exhibitions

In 2010 the *Articulated Head* appeared in two exhibitions [26,27], both connected to scholarly conferences but open to the general public. In the same year, the *Articulated Head* was a finalist for an Australian engineering award [28] and was selected to be displayed throughout 2011 in the Powerhouse Museum, Sydney, Australia (this was subsequently extended for another year) [29]. The Powerhouse Museum is visited by approximately 480,000 visitors per year [30] and with the *Articulated Head* located not far from the main entrance, most visitors encounter it at least briefly. Thousands of interactions between the audience and the *Articulated Head* have been observed and some recorded. They last anywhere from only a few seconds to more than half an hour. The general pattern is that the appearance of the installation itself (a "strange"-looking robot within an enclosure) attracts the attention of visitors from afar, the robot movements fuel curiosity on approach and kick-start the interaction, until finally the language-based communication with the integrated chatbot becomes the primary center of the interaction. Visitors notice at various times that the *Articulated Head* attempts to mimic head poses, and participants start to play with it.

Among the interactions observed on several occasions were also games similar to hide-and-seek played by small children with the *Articulated Head*. These games turned out to be remarkably successful: The children waited until they were tracked—that is, the head was looking at them—and then ran to a new location right at the enclosure, trying to hide behind the wooden support for the glass barrier or behind the information kiosk (see Fig. 4 for the spatial layout). The *Articulated Head* uses the monovision camera mounted on the monitor for face detection, but the presence of people is detected with a static stereo camera mounted above its enclosure. Thus, it does not need to orient toward a person and have an unobstructed line of sight to register the person. However, the people-tracking software requires a minimum height threshold for tracking. It was set to 0.5 m. Thus, the children could hide in the tracking shadow simply by crouching, but when peeking above their assumed hiding barrier they returned into the tracked area, and the *Articulated Head* oriented its head toward them, including, of course, adjusting its elevation angle. Therefore, it would appear to look down at them after having rediscovered them when they were carefully—but not cautiously enough—peering from their hideout. Following their discovery, the children would quickly run to a new location and hide again.

This is in our view a strong demonstration of evoked agency. Although children may attribute agency to many objects (e.g. dolls and stuffed animals) and the displayed face on the monitor of the *Articulated Head* most likely played a role as well, movements appearing not to be related to the actions of the children would destroy the perception of agency (and the game): The correct sequencing and timing of the robot movement is crucial. The properties of the interaction have to fulfill certain constraints (see, for example, Terada et al. [31]), and current research is only scratching the surface of what precisely these constraints are. However, more important for our work with respect to the control system of the *Articulated Head* is the fact that the game exemplified emerging behavior, since we never planned for a game like this to be played by the *Articulated Head*. It demonstrates human-machine interaction emerging from situational context and predispositions for social interaction—grounded sometimes in remarkably simple principles.

Acknowledgments

The authors would like to thank Zhengzi Zhang for his inter-component communication software and five anonymous *Leonardo* reviewers for their insightful comments. We wish to acknowledge the support of NHMRC/ARC grant TS0669874.

References and Notes

Unedited references as provided by the authors.

1. A. Clark. *Being there: Putting brain, body, and world together again.* The MIT Press, Cambridge, USA, 1997.

2. C. Kroos, D.C. Herath, and Stelarc. "The Articulated Head: An intelligent interactive agent as an artistic installation." In *Proceedings of International Conference on Intelligent Robots and Systems*, St. Louis, MO, USA, 2009.

3. C. Kroos, D.C. Herath, and Stelarc. "The Articulated Head pays attention." In *Proceedings of International Conference on Human-Robot Interaction*, pp. 357–358, Osaka, Japan, 2010.

4. J. Gibson. *The ecological approach to visual perception.* Houghton Mifflin, New Jersey, USA, 1979.

5. B. Hommel, J. Musseler, G. Aschersleben and W.

Prinz. "The theory of event coding (TEC): A framework for perception and action planning." *Behavioral and Brain Sciences*, 24(5):849–878, 2001.

6. The *Thinking Head* project was funded jointly by the Australian Research Council and the National Health and Medical Research Council.

7. M. Mori. "The uncanny valley," *Energy*, 7(4):33–35, 1970.

8. R.S. Wallace. "The anatomy of A.L.I.C.E.," in R. Epstein, G. Roberts & G. Beber, editors, *Parsing the Turing Test*, pp. 181–210. Springer Netherlands, 2009.

9. A. Waytz, K. Gray, N. Epley, and D.M. Wegner. "Causes and consequences of mind perception." *Trends in Cognitive Sciences,* 14(8):383–388, 2010.

10. P. Cavanagh. "Attention routines and the architecture of selection." In Michael I. Posner, editor, *Cognitive Neuroscience of Attention,* pp. 13–18. Guilford Press, New York, 2004.

11. D. Heinke and G.W. Humphreys. "Computational models of visual selective attention: A review." In G. Houghton, editor, *Connectionist Models in Psychology.* Psychology Press, Hobe, UK, 2004.

12. S. Ohayon, W. Harmening, H. Wagner and E. Rivlin. "Through a barn owl's eyes: Interactions between scene content and visual attention." *Biological Cybernetics*, 98:115–132, 2008.

13. R.J. Peters and L. Itti. "Computational mechanisms for gaze direction in interactive visual environments." In *Proceedings of 2006 Symposium on Eye Tracking Research & Applications*, San Diego, California, USA, 2006.

14. T. Bosse, P.-P. van Maanen and J. Treur. "A cognitive model for visual attention and its application." In *Proceedings of International Conference on Intelligent Agent Technology*, pp. 255–262, Hong Kong, 2006.

15. Y. Sun, B. Fisher, H. Wang and M. Gomes. "A computer vision model for visual-object-based attention and eye movements." *Computer vision and image understanding*, 2008.

16. Y. Kim, R.W. Hill and D.R. Traum. "A computational model of dynamic perceptual attention for virtual humans." In *14th Conference on Behavior Representation in Modeling and Simulation (brims),* Universal City, CA., USA, 2005.

17. J.A. Driscoll, R.A. Peters and K.R. Cave. "A visual attention network for a humanoid robot." In *Proceedings of Intelligent Robots and Systems*, Vol. 3, pp. 1968–1974, 1998.

18. O. Déniz, M. Castrillión, J. Lorenzo, M. Hernández and J. Méndez. "Multimodal attention system for an interactive robot." In *Pattern Recognition and Image Analysis*, pp. 212–220. Springer, Berlin, 2003.

19. J. Morén, A. Ude, A. Koene and G. Cheng. "Biologically based top-down attention modulation for humanoid interactions." *International Journal of Humanoid Robotics*, 5(1):3–24, 2008.

20. P. Bachiller, P. Bustos and L.J. Manso. "Attentional selection for action in mobile robots." In *Advances in Robotics, Automation and Control*, pp. 111–136. InTech, 2008.

21. C. Breazeal and B. Scassellati. "A context-dependent attention system for a social robot." In *Proceedings of the 16th International Joint Conference on Artificial Intelligence*—Vol. 2, pp. 1146–1151, San Francisco, CA, USA, 1999.

22. C. Kroos, D.C. Herath and Stelarc. "From robot arm to intentional agent: The Articulated Head." In Satoru Goto, editor, *Advances in Robotics, Automation and Control,* pp. 215–240. InTech, 2011.

23. The current exhibition is at the Powerhouse Museum, Sydney, Australia.

24. R.S. Anderssen, S.A. Husain and R.J. Loy. The Kohlrausch function: Properties and applications. In *Proceedings of 11th Computational Techniques and Applications Conference,* Vol. 45, pp. C800–C816, 2004.

25. See review in Heinke and Humphreys [11].

26. NIME++ (New Interfaces for Musical Expression), 15–18 June 2010, Auditorium, University of Technology Sydney, Australia.

27. SEAM: Agency & Action, 15–16 October 2010, Seymour Centre, University of Sydney, Australia.

28. The *Articulated Head* was a finalist in the Engineering Excellence Awards (Sydney section of Engineers Australia).

29. Engineering Excellence Awards, 2011, Powerhouse Museum, Sydney, Australia.

30. <www.powerhousemuseum.com/about/aboutFacts.php>; accessed 24 January 2012.

31. K. Terada, T. Shamoto, Haiying Mei and A. Ito. "Reactive movements of non-humanoid robots cause intention attribution in humans." In *Intelligent Robots and Systems,* 2007. IROS 2007, pp. 3715–3720.

Manuscript received 15 November 2010.

Christian Kroos received his M.A. and Ph.D. in Phonetics and Theatre Studies from the Ludwigs-Maximilians-Universität, Munich, Germany. He has conducted interdisciplinary research covering computer vision, cognitive sciences and robotics at the Institute of Phonetics and Speech Processing at Ludwigs-Maximilians-Universität (Germany), ATR International (Japan) and Haskins Laboratories (U.S.A.). Besides his interest in robotic agents, he is still fascinated by human speech production and the evolution of language.

Damith Herath received his Ph.D. in Robotics from the University of Technology, Sydney, in 2008 and a BSc (Hons) in Production Engineering, University of Peradeniya, Sri Lanka, in 2001. He held a doctoral fellowship at CAS prior to joining MARCS Institute on the Thinking Head Project as the Research Engineer. Currently he leads several robotic projects that explore human-robot interaction (including reciprocal influences between the arts and robotics).

Stelarc's projects explore alternate anatomical architectures. He has performed with a Third Hand, a virtual body and a 6-legged walking robot. An ear that will be Internet-enabled is being surgically constructed and cell-grown on his arm. In 1997 he was appointed Honorary Professor of Art and Robotics at Carnegie Mellon University. In 2003 he was awarded an Honorary Degree of Laws by Monash University. In 2010 he was awarded the Prix Ars Electronica Hybrid Arts Prize. He is currently Chair in Performance Art, School of Arts, Brunel University. Stelarc's artwork is represented by the Scott Livesey Galleries in Melbourne.