netzspannung.org
Wissensraum für digitale Kunst und Kultur

SENGERS, Phoebe
LIESENDAHL, Rainer
MAGAR, Werner
SEIBERT, Christoph
MÜLLER, Boris
JOACHIMS, Thorsten
GENG, Weidong

The Enigmatics of Affect

Published on netzspannung.org:
http://netzspannung.org/about/mars/projects/pdf/influencing-machine-2002-1-en.pdf
2 February 2005

First published: Accepted to the Designing Interactive Systems Conference, London, June 2002.

ACM COPYRIGHT NOTICE. Copyright © 2002 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Publications Dept., ACM, Inc., fax +1 (212) 869-0481, or permissions@acm.org.

IMK
Fraunhofer Institut Medienkommunikation

The Exploratory Media Lab
MARS Media Arts & Research Studies

The Enigmatics of Affect

Phoebe Sengers[+], Rainer Liesendahl, Werner Magar, Christoph Seibert, Boris Müller[–], Thorsten Joachims[⊥], and Weidong Geng
MARS – Exploratory Media Lab
Fraunhofer Institut für Medienkommunikation

Pia Mårtensson and Kristina Höök
Swedish Institute of Computer Science

[+] Now at Computing and Information Sciences / Science & Technology Studies, Cornell University, sengers@cs.cornell.edu.

[–] Now at esono.com, boris@esono.com.

[⊥] Was at Knowledge Discovery, Fraunhofer Institut für Autonome Intelligente Systeme; now at Computer Science, Cornell University, tj@cs.cornell.edu

ABSTRACT
Affective computation generally focuses on the informatics of affect: structuring, formalizing, and representing emotion as informational units. We propose instead an enigmatics of affect, a critical technical practice that respects the rich and undefinable complexities of human affective experience. Our interactive installation, the Influencing Machine, allows users to explore a dynamic landscape of emotionally expressive sound and child-like drawings, using a tangible, intuitive input device that supports open-ended engagement. The Influencing Machine bridges the subjective experience of the user and the necessary objective rationality of the underlying code. It functions as a cultural probe, reflecting and challenging users to reflect on the cultural meaning of affective computation.

Keywords
affective computation, affective interaction, believable agents, agent architecture, interactive installations, interactive sound, tangible interfaces, interaction design, critical technical practices

INTRODUCTION

> *"Das ganze ist eine Allegorie – eine fortgeführte Metapher!" [18]*

Affective computation, or the development of computational systems which can be aware of and respond to human emotions, has become the focus of a great deal of attention in the Artificial Intelligence (AI) community. These approaches to affective computation often focus on what one might call an informatics of affect, in which emotions are treated as units of information. Emotions are analysed, classified, discretised, and formulated as units whose purpose is to inform cognition or be communicated. The often-used integrative cognitive theory of emotion of Ortony, Collins and Clore, for example, defines emotions in terms of a set of discrete, basic types and focuses on the cognitions or reasoning which may give rise to them [12]. Once a set of emotional units is defined, input devices can be designed which can turn physiological responses into information. For example, Fernandez, Scheirer, and Picard's Expression Glasses measure the movement of facial muscles and classify the resulting expression into a small, discrete set of emotions [15]. Ark's Emotion Mouse extends a normal computer mouse to deduce users' emotional states from physiological information such as pulse and galvanic skin response[2].

Frequently in this tradition, emotions are subsumed to rationality or effectiveness. Damasio's influential arguments for the importance of emotion in scientific research, for example, gain currency from the idea that emotion is necessary for true rational behaviour [13]. Similarly, Picard's ground-breaking work on Affective Computing argues that computers must be able to process emotion in order to function maximally effectively with human beings[30].

While defining, classifying, creating logical structure for, and understanding the relationship of rationality to emotions can be useful exercises, bringing with them the pleasures of both computability and scientific respectability, we believe this mindset is in danger of missing a fundamental point: affect is not just a formal, computational construct, but also a human, rich, enigmatic, complex, and ill-defined experience. Rationalizing it may be necessary to make it computable, but an affective computation that truly inspires and incorporates human emotion must include a broader cultural perspective, in which the elusive and nonrational character of emotion does not need to be explained away. From this perspective, computation may be used, not to acquire and reason about user's emotional states, but rather to create intuitive experiences of affect by the user during interaction. We term this alternative approach to affective computation an *enigmatics* of affect.

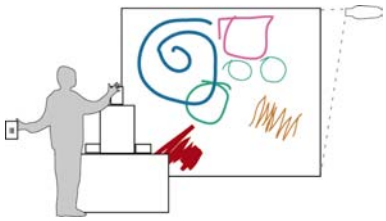
This approach does not, however, liberate us from the constraints of computability. In order to function, technology requires us to discretize, analyse, and formalise. A substantial design challenge in constructing a technical system that supports enigmatic, open interpretation, then, is the need to bridge the rational objectivity of the soft- and hardware with the interpretational complexity of users' subjective experiences. Doing this well requires combining technology design in the tradition of AI research with cultural analysis of the meaning and metaphors embedded in the interaction, in the spirit of a critical technical practice [1].

Here, we will describe the Influencing Machine, an interactive installation we built to explore issues in the enigmatics of affect. In this installation, users influence the emotions of an (invisible) artificial agent, which expresses its emotions by generating real-time, dynamic, child-like scribblings and through an emotionally evocative soundscape. For this system, the underlying technology that the system requires to process emotions and generate expressive behaviour is in the background, with the focus on users having experiences with the system that engage their emotions and critical and interpretive faculties. Users should ask, "Which emotions do I see? Do I agree with the machine? Do I believe these drawings express emotions? What is my relationship to this emotional machine?" In order to support this questioning, the input and output of the system are deliberately left complex, enigmatic and open to interpretation.

The Influencing Machine functions as a boundary object, floating between the worlds of algorithms and interpretation, between technology design and cultural analysis. It forms itself a cultural commentary on the nature of emotion in the culture of computation, allowing users to take part by experimenting with and reflecting upon their experiences with the machine.

CONCEPT

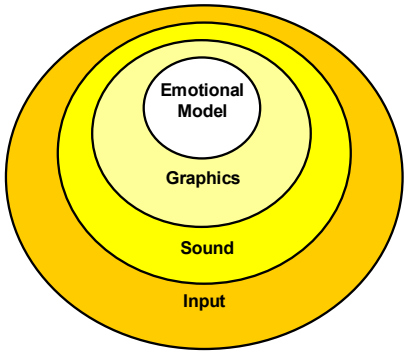
Two people enter a small room. Child-like scribbling appears across a wall: jagged lines, circles, spirals, and other shapes build up, overlap, fade away. Scattered throughout the room are postcards with art prints; on a table stands a wooden mailbox. One person picks up a card and tentatively puts it in the box. Unusual and musical sounds begin to play. Drawings change speed, color, pressure, form. The people begin sorting through cards, dropping them in the box and seeing how the graphics and sound change. They play, experiment, discuss: "How is this reacting to us?" "How do you think this works?"



Setup of the Influencing Machine

Technically, the system works by using the input postcards to influence an internal emotional model. These internal emotions trigger sounds and the selection of drawing behaviours and their dynamic parameters: speed, colour, size, pressure, etc.

While this technical description is precise and clean, the emotional interpretation of the graphical output and postcards by users is complex, incompletely specifiable, open-ended, and strongly culturally influenced. Rather than forcing users to interact with the system using its internal, clean, unnatural emotional representations, in the design of the Influencing Machine we wanted to enhance users' intuitive, interpretive experience of the system. The basic problem for this design is the disjunction between the internal representation of discrete emotions that drive graphics and sound and the desired complex, open-ended interpretations of the user



The Influencing Machine technology is organized into layers of cultural complexification, surrounding the core emotional model.

Conceptually, the system is organized into a set of concentric layers, which start out as highly formal, clean and technical, and move out to incorporate more and more complex and ill-defined aspects of user experience. The central layer, the emotional model, forms the rationalized core of the system. To this we added first algorithms for graphical output, which are both technically substantial and begin to evoke emotional interpretation in complex, subtle, and multivalent ways. The design of the input system moves out further, being based on user interpretation of emotion and bringing this into relationship to the internal emotional model and thereby to the graphics. Finally, we added an intermediate layer of sound, which mediates and further expresses the relationships between the postcard input and the complex output. Next, we will describe the design motivation and technology development for each of these components in turn.

EMOTIONAL MODEL

The internal emotional model is the cornerstone of the system; the other modules function by referring to and changing the emotional model. Internally, emotions are represented using a discrete set of 22 emotions, which are organized as 11 pairs of opposite emotions: anger / peace, transcendent / physical, insecure / assured, primal unity with the world / sense of self, control-stasis / flow-dynamics, warm / cold, happy / sad, aggressive / passive, introverted / extraverted, safe / fearful, energetic / weak. We chose this set of emotions by analysing the emotional terms used in the depth-psychological analysis of children's drawings [31] [12], which corresponds to the sorts of emotions we expected to be able to express through the graphical output of the machine.

These emotions did not correspond well to commonly available cognitive models of emotion, and we were forced to develop our own model. Our focus is not on developing a complex theory of and structure for the internal emotions, but on how emotions affect the system's behaviour and thereby user interpretation. Hence, we implemented the emotional model as a minimum-commitment model

in which an emotion simply maintains a value between 1 and 10. Opposite emotions are correlated; i.e., when the agent is happy, it cannot also be sad. Emotional levels decay over time, so that, if left alone, the agent tends to a neutral emotional state.

GRAPHICS

The emotional model is a purely formal, informational construct. The system graphics must begin to translate that clean model into visible forms and dynamics that evoke more open-ended emotional meaning for users. The graphical output of the system should hint at a subjective view of the world. It should be enigmatic but invite interpretation.

This goal is contrary to much contemporary computer graphics research, which is focused on the generation of objective, photorealistic representations of 3D models. Even nonphotorealistic rendering, which creates renderings of 3D models in an artistic style, does so in a directly representative way, simply showing the objects using a painterly rather than photographic style. We needed to look elsewhere.

Our inspiration [1]

We found our inspiration from an opposite practice, children's drawings. While often failing entirely at realistic representation, these drawings communicate a great deal of charm and personality. We wanted to generate similar imagery automatically. Like children's drawing, this generation must be dynamic, i.e. an observable process in real-time, with emotions being reflected not only in the end-result – which is often in children's drawing beside the point - but also in the dynamic qualities of the strokes that make up the picture. We wanted to draw freely, without the input of 3D models and representations. And we hoped the drawings would be charming.

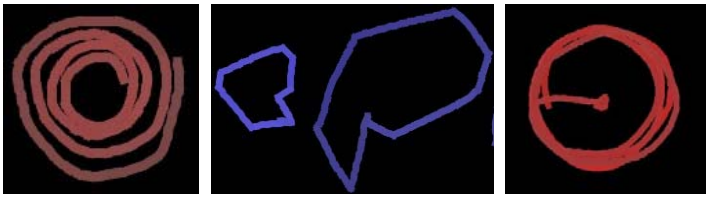
There are two major drawing programs which act as precedent for our work in this area. The first is Harold Cohen's Aaron, a program which generates drawings by controlling a robot arm [12] [24]. Aaron models an expert artist (i.e. Cohen himself!), and can generate a wide variety of drawings of natural scenes with human figures in Cohen's (now Aaron's) signature style. We are interested in less expert graphical behaviour. Our most direct inspiration is therefore Ed Burton's Rose [10][11], a program that generates childlike drawings of 3D models. The Rose program takes as input a CAD/CAM model, analyses it in a way metaphorically similar to children's perception, and then produces a childlike drawing of it. We add to Burton's work by including the ability to model and communicate emotion through drawings; by generating drawings dynamically, i.e. where not only the output, but also the real-time process of doing the drawing is child-like; and by eliminating the input model, which is not realistic for children's art before the age of 4 or 5, and instead driving the 'content' of the drawing from the agent's emotional state.

Ed Burton argues strongly that his work should not be seen as an attempt to directly model children's drawings within the computer. Instead, he explores the kinds of systems one could build when drawing on particular metaphorical understandings of children's drawings. We follow this view by being inspired in the construction of the system by stories of children's drawings, and especially stories of their emotional meaning. We draw these stories largely from the depth psychology literature, which unlike the cognitive science literature discusses not the general properties of drawing, but the individual emotional and subjective meaning of these drawings in a narrative form [35][31].

The use of these depth-psychological theories is problematic from a scientific perspective. They are no longer accepted by art pedagogians because they rely too heavily on psychologists' interpretations. But because our goal is not to anthropomorphically recreate children's drawing processes in the computer, but to generate child-like, emotionally evocative drawings, heavily interpretational theories, however incorrect, are most appropriate for us.

Implementation

The fundamental design issue in the Influencing Machine is the need to mediate between the internal, simple emotional model and the complex interpretations of the user. For the graphics of the system, this means the drawings should not be a simple, one-to-one mapping of internal emotion to an observable graphical token (e.g. a 3-inch red triangle that appears whenever the agent is angry). Instead, graphics should be dynamically, be newly generated in real time, and evoke complex and open-ended interpretations of multiple simultaneous emotional states.



Examples of generated scribblings

The basis of the graphics engine is a behaviour-based, reactive control architecture similar to those used to control animated computer characters. This architecture selects drawing forms and

[1] This image is a detail from Family with Popsicles, by Amy Morrisett.

styles based on the current emotional state, and executes them by sending commands to control the movement of a virtual pen over the screen. These commands are then implemented by a separate stroke renderer, which implements a simple model of the natural variation of human drawing. Here, I describe each of these systems in turn.

BEHAVIOUR ARCHITECTURE

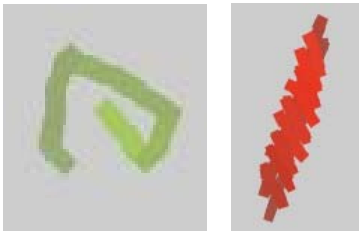
The graphics engine must generate drawings in real time that reflect in content, form, and dynamics the current state of the emotional model, which may include up to 11 different emotional states to express. Choosing what to draw and how to draw it is a variation of one of the fundamental problems in agent research: the appropriate, dynamic selection and execution of activity (see, for example, [23]). Hence, we solve the drawing problem by developing a specialized agent architecture. At the lowest level, the architecture must continually redecide which low-level attributes (color, pressure, wobbliness or noise, and speed) best express the current emotion set. At a higher level, it must decide on and execute the form that would best express the current emotions. The agent architecture must allow the agent to make complex and interdependent decisions about what to draw, while remaining reactive to new emotional input from users. It must steer the drawing process over time. Most importantly, it must incorporate multiple emotional factors into every aspect of its decision-making.

These requirements suggest the use of a multiple-level reactive control architecture, similar to Brooks's subsumption architecture [7][9] or Airsoar [28]. In such an approach, the different levels of the architecture operate continuously and in parallel, each continually monitoring environmental conditions and changing its actions based on changes to the environment. Our architecture has two levels: (1) a low level, which sets the drawing attributes, and (2) a high level, which selects and executes the drawing forms. Both levels incorporate multiple emotions in every aspect of decision-making and execution. In order to avoid watering down the emotional effects by averaging between all possible emotions, the behaviour architecture considers only the top 4 strongest emotions in its decisions.

Low level Architecture

Scribbling with low (left) versus high (right) pressure.

The low level resets the drawing attributes – pressure, speed, wobbliness, and colour – once per animation frame (15 times a second) to reflect the current emotional state. For pressure, speed, and wobbliness, this is done by assigning a preferred level for each emotion that can be expressed through that attribute. For example, extreme anger will greatly increase speed, while extreme sadness will greatly reduce it. At any point in time, the current value of the attribute is calculated by multiplying the preferred level for each emotion by the current strength of that emotion (measured from 0.0 to 1.0) and averaged over the strengths of the total number of emotions considered. Thus, the strongest emotions have the most influence on the resulting level of the attribute.



Color variation in and between emotions; introverted (left) vs. aggressive (right) coloration

Colour is calculated in a different manner, since colour cannot meaningfully be simply averaged. Instead, each emotion that is expressible through colour is assigned a particular colour range. This colour range consists of 3 subranges of hue, saturation, and brightness. The ranges are marked with how much they express the corresponding emotion; so, for example, a less saturated blue will express more coldness than a more saturated blue. In this way, the colour can be varied depending on the level of the emotion being expressed, so that subtleties of emotional strength can be conveyed through shades of colour. When the agent begins to draw a new form (i.e., has lifted its pen from the paper), it selects the colour that corresponds to the strongest emotion expressible through colour.[2] While drawing a form, the agent varies the saturation and brightness of that colour to express the current emotional strength. In addition, a small amount of noise is added to the saturation and brightness of the colour so that the stroke can still be seen when the agent is scribbling over previously drawn lines. This also adds to the colour interest and attractiveness of the drawn scribblings.

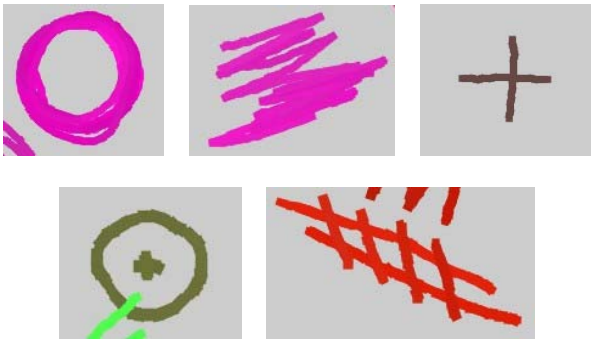
High Level Architecture

The high-level architecture must continually redecide the form to draw and execute it appropriately for the currently active emotions. We implement this in a behaviour-based framework, as described e.g. in [8]). In this framework, each form the agent can draw is implemented as a 'behaviour,' a self-contained piece of code which is reactive to changes in the environment – in this case, the emotional model – and responds to them by executing actions – in this case, by sending drawing commands to the pen. The agent stays up to date by using an action-selection algorithm [22]; once per frame, it redecides whether the current form is appropriate, and then engages in a small amount of computation for that form. This allows it to continuously respond to changes in input from the emotional model.

The forms that have been implemented are those that have been identified in the literature on the early stages of children's drawings (see especially [19]). It is important to note that a single behaviour does not correspond to a single graphically visible form. Rather, a

[2] If there are no strong emotions, it uses beige.

behaviour describes procedurally how a shape is drawn, and can result in many graphical variations of that form. This is essential for the interpretational complexity of output (i.e., the forms are not simply ever-repeating tokens). It is also essential for expressing emotion; each behaviour encodes within it the variations necessary to express many different emotions by varying the appearance of the form.



Some of the available forms: circular scribbles, line scribbles, crosses, centred circles, ladders.

A single behaviour can draw many variations of a form. These are all line scribbles.

Technically, a behaviour is implemented as a Java class with two components: activation and appropriateness. An **activation** level states whether the behaviour is available to be executed. Behaviours are organized developmentally; initially the agent will only draw simple forms, gradually adding to its repertoire over the course of the interaction. A behaviour is activated when its developmental state has been triggered. This allows it to begin being considered for execution.

The behaviour's **appropriateness** calculator determines how appropriate the behaviour is for expressing the current emotional state. The degree to which any particular form expresses an emotion is derived in part from the depth-psychological literature on children's drawings, which states the emotional interpretation of different shapes. To this is added considerations about the concrete implementation of the behaviour and how it can vary with different emotional states. The appropriateness calculator is used by the behaviour architecture to select the behaviour that can best express the current emotional state. If a behaviour is not activated, its appropriateness is zero.

Once a behaviour class has been chosen for execution, it is instantiated to create a concrete, running behaviour. These behaviour instances have the following components. A set of **parameters** specify the detailed functioning of this instance of the behaviour class. Parameters vary depending on the form, and include such things as size, radius, or number of iterations. The parameters can be modified externally to alter the functioning of the behaviour, even while the behaviour is running. Communication between behaviours occurs when they modify each other's parameters.

Some of the behaviour's parameters are particularly useful for expressing emotions. These are defined as **emotional variables**. Emotional variables recalculate their values every time they are used in order to reflect the current emotional state. This is done by first setting a "home value," which represents the variable's value in the absence of emotional influences. Coefficients are set up that express how much this home value should be varied based on emotional state. For example, when introverted, the radius of a circle should be drawn smaller; this is represented by choosing a coefficient less than 1 for the introverted emotional state. When the value of an emotional variable is retrieved, it calculates the combined emotional influence of the current emotional state on the variable and returns the "home value" of the variable multiplied by that coefficient. This allows the parameters of the behaviour to adapt continuously to the changing emotional state.

A behaviour has a run method that, when executed, implements one step of the behaviour. A "step" should be brief, no more than 100 milliseconds. This is to allow real-time response, by rethinking the choice of what to do frequently between executions. The run method is passed the current drawing attributes from the low level of the architecture, which it uses to influence the path and pressure of the stroke commands it uses. Typically, this method engages in some computation to determine the next stage of the stroke, and then implements this stage by calling rendering methods of the stroke system. Frequently, behaviours also create subbehaviours, modify their parameters, and run them in order to implement their own functionality. For example, the centred circle spawns first a circle behaviour, then a cross behaviour, where the parameters of the cross are altered to be at the centre of the previously drawn circle.

Finally, a behaviour maintains a **drawing sensor**. This is a data structure that keeps tracks of the points that this behaviour has drawn so far. These points are returned by the stroke system that implements drawing commands (described below), and represent the actual points drawn - rather than the points where the behaviour attempted to draw; because of the noise in the drawing system, which simulates human-like drawing, there is usually a divergence between these two! These sensed points can then be used to modify later behaviour. For example, the "ladder" drawing behaviour draws two lines, then draws crossbars between the sensed points from the lines.

The behaviour architecture contains a set of **behaviours**, as defined above. It maintains a currently running **behaviour instance**, i.e. the behaviour which is currently in the middle of execution. A **selection** method allows the architecture to choose among all possible behaviours The selection method works by checking the appropriateness of each behaviour for the current emotional state, then choosing probabilistically among them, weighted by the

appropriatenesses[3]. The behaviour architecture is called once per animation frame. It first runs the selection method to see which behaviour is the most appropriate for running based on the momentary emotional levels. If this is not the currently running behaviour, it ends that behaviour and starts the more appropriate one. It then uses that behaviour's run method to run one step of the behaviour.

In practice, 21 behaviours were implemented. Writing a behaviour involved the most work in two areas: (1) fine-tuning the geometry and pacing of the behaviour; (2) varying the behaviour properly for the 22 different emotions, including setting the appropriateness of the behaviour and varying the behaviour's parameters. The first behaviours written generally dealt with pure forms --- drawing lines, drawing circles; then came behaviours which varied forms --- for example, varying a circular scribble to create a spiral; finally came behaviours which combined already existing behaviours --- for example, a circle with a cross at the centre. These combination behaviours were trivial to write, implying that further extension of the behaviour to deal with complex combinations of forms should be straightforward.

STROKE RENDERING

The behaviour architecture selects commands for the agent's virtual pen, including the following: setting the colour; setting the pressure; resetting the position of the pen; and moving the pen towards a point. The resulting strokes create shapes on the screen. While the behavioural drawing system can suggest strokes that form, for example, a circle, the output will not seem human-like if the strokes move directly, cleanly, and exactly along the line which the behavioural drawing system suggests. Instead, the line must be, like those in children's drawings, wavering, meandering across the page with error and noise, and with the agent's ability to control the resulting line decreasing with the speed at which the line is drawn. It is the responsibility of the stroke rendering subcomponent to determine a suitably errorful path for the stroke to create the impression of liveliness in the resulting line.

The stroke renderer varies the graphic output in the following two ways. First, it has a simple model of an arm with momentum. When the system has calculated the direction it wants to go in, it must work against its previous momentum to go in that direction, by using a variable Force, specified by the calling system. By making Force low, the resulting strokes curve more, whereas with a high Force, the resulting strokes can have sharper angles. This adds a sense of dynamics to the resulting drawings. The second algorithm the stroke renderer uses is to add some noise to the movement. When a stroke command comes in, the system does not go directly to the target point, but towards a noisy approximation of it. As the system comes closer to the target, it chooses new noisy approximations of its target. In this way the system meanders more or less uncertainly towards the target, creating the effect of lower stages of motor control as one might expect in a small child.

Arm simulation adds a sense of dynamics to the movement of the drawings.

The stroke itself must also appear to be something drawn by an instrument a human being would typically use, varying, for example, in colour and pressure. Currently, only the stroke pressure is varied. More sophisticated stroke forms could be generated by selecting a set of colourful painted images as reference points, and pre-building their non-photorealistic rendering models, which include the composition of painting genres, colour, size, lighting, shadow, texture, background etc. During the real-time drawing of graphics, the parameters of strokes could be automatically computed by these non-photorealistic rendering models, expanding the range of human-like strokes generated.

INPUT

The graphical engine of the Influencing Machine moves out from the straightforward definitions of the internal emotional model to create dynamic imagery that invites interpretation. This imagery is, however, driven by the internal emotional model, rather than by the emotional interpretation of the user. In order to connect the graphics more directly to the complexities of user experience, we need to design an appropriate input device.

The input device for the Influencing Machine must allow users to influence the behaviour of the system through its internal emotional values while enhancing users' interpretational experience of the machine and enabling critical engagement. A typical affective computation interface might try to measure and represent users' unconscious emotional state; this, however, would not support critical engagement, since users may not be aware of how they are interpreted. It would also make multi-user engagement, which supports critical engagement by providing the opportunity for conversation, difficult. On the other hand, a screen interface with sliders would give users precise control of internal emotional values. But this, again, does not enhance the experience of the process: it instead forces users to interact with the system using its internal representation, not using the concepts the user is likely to come up with spontaneously in observing the system's output.

[3] The selection method incorporates the following techniques to handle the problems of dithering and inappropriate persistence as identified by Blumberg [5][6]. The appropriateness of the currently running behavior is doubled, in order to reduce the problem of dithering. Behaviors also inhibit themselves: when a behavior has run, it triggers an increase in its inhibition level, which in turn decreases its appropriateness. This inhibition is cumulative; once a behavior has run a few times it is unlikely to be chosen again. This increases the variance of observed behavior for users.

An additional complication is raised by the unclear ontological nature of the user's relationship to an emotional machine. Culturally, the notion of an emotional machine is paradoxical; an emotional being is something we relate to, care for, and nurture, while a machine is something we instrumentally use and control. How should humans relate to emotional machines? How should emotional machines in turn understand their relationships to humans?

One answer to these questions inspired by the metaphor of the influencing machine, a paranoid delusion first described by Victor Tausk [36] and extensively described by Bruno Bettelheim in his case study of Joey, a boy who believed he was mechanical [4]. People suffering from the influencing machine delusion feel that they are being controlled by a machine that projects hallucinations, produces or removes thoughts, feelings, and physical sensations, and changes one's bodily composition. This influencing machine is a projection of part of a person's sense of self; i.e. his or her sense of self is split into two parts that share control of the person.

In this way, the influencing machine is a model of what it is like to share control between two entities, a concept that maps to the user-emotional agent relationship. It is easy to imagine that an emotional machine could experience the user as a kind of influencing machine, who is not directly accessible to the machine but causes changes to the machine's actions, thoughts, and emotions. In the Influencing Machine system, the user should not directly control but indirectly influence the system's behaviour, acting as the influencing machine for the drawing agent.

Design

These observations lead to the following design constraints for input. Input, like perceived output, should be fuzzy, supporting open-ended engagement with the system without imposing a particular, limited set of emotional interpretations. It must allow this rich, analogue interaction while supporting discrete internal processing. The input device should be easy to understand and enjoyable. It should support multiple simultaneous users, who can share their interpretations, reflections, and criticisms. Users must be able to influence, but not control, the agent - and to know they can and how.

To fulfil these requirements, we decided to create a tangible input device employing open-ended, imprecise representations. When users enter the Influencing Machine installation, they find physical postcards with emotionally evocative art prints scattered around the room. On a table stands a wooden "mailbox" or "voting box", into which users can put the postcards, thereby sending emotional messages to the agent. These messages alter the current emotional state of the agent, which in turn alters the soundscape and the form and dynamics of the graphics that are being seen by users.

From a technical perspective, each postcard represents an emotional profile, or a set of emotional values. This profile is designed by hand to match the emotions that the image on the postcard may evoke. It is important to note that input emotions do not directly set the agent's emotions, but are added to or subtracted from the agent's current emotional state. Thus, a highly introverted agent cannot be turned into an extraverted agent with one postcard; instead, extraverted postcards must be used multiple times, with the agent's emotional state gradually becoming more extraverted with each card.

From a human perspective, users pick up postcards and interpret the emotions they contain. By sending a card, users send their personal, emotional reactions to the system, hoping to evoke the same emotions in the agent, or maybe with only a vague idea of the link between card and response. Observing the resulting drawings and sounds, users consider their opinion of the card, comparing it to and often commenting on their interpretation of the agent's behaviour.

The tangibility of the interface brings the experience outside of the machinic and into the physical world of users, who use the device together. It looks whimsical and approachable. Appropriately chosen art postcards cry out for emotional interpretation; they call on users' natural intuition.[4] By using sets of emotions with different levels for each card, we avoid a push-button, one-to-one interface, and have the option for nearly arbitrarily many different input objects with different effects. Mapping cards to sets of emotion values is a straightforward way to make the emotional complexity of the postcard imagery computable.

Implementation

For the Influencing Machine, we selected a set of about 40 postcards from art history with art historian Gabriele Blome. Postcards were selected according to the following criteria: clear representation of more than one emotion; not evoking too many emotions; only evoking emotions corresponding to the internal emotional model; and broad coverage of the emotional model. The postcards were marked on the backs with UPC codes and laminated.

Some of the art postcards used for the Influencing Machine.

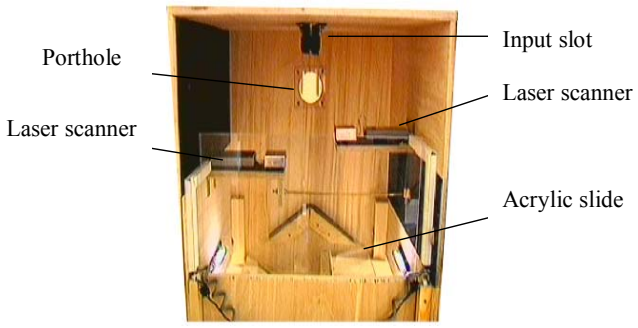
The input device itself, the Mailbox, is a wooden box of 45x45x60cm with a slot in the top. The rectangular shape of the slot constrains the

[4] Concerned about the potential arbitrariness of interpretation of art prints, we also developed a set of input postcards with one to three solid colors, based on Eva Heller's empirical theory of the emotional meaning of colors [17]. In practice, users found the precisely defined color postcards much harder to interpret than the less precise art postcards.

axis of orientation of the postcards deposited through it. The box contains two extendable drawers, one on each side. Two portholes on the front and back with movable shutters offer an inside view of the workings for those curious about what happens inside.

The constructed Mailbox.

Inside the box, two laser scanners are installed opposite each other, which capture the barcode on the postcard regardless of the postcard's orientation. The laser scanners are offset from each other, so that they do not interfere with each other's readings. After being thrown in, the cards land upon a small metallic tube that stops their fall long enough to ensure that they are read. The cards then fall onto an acrylic slide and into one of the extended drawers, where they are again available for use.



The internal workings of the Mailbox.

SOUND

By themselves, the emotional evocations of the graphics of the system are complex and multivalent, and hence difficult to interpret. The art postcards, too, are richly evocative – and thereby not clearly and straightforwardly classifiable. There is a danger of the user being, not stimulated to autonomous interpretation, but simply confused by the complexities of the input and output of the system.

Because of this difficulty, we decided to add an extra layer of expression through an evocative soundscape between the input and graphical output of the system. These sounds play in the background, providing additional support for disambiguating the potential emotional implications of the dynamic graphics. They also provide additional feedback for input to the system (answering the question, "did that do anything?"), by responding directly and immediately to the emotional implications of the input postcards. In our experience, they also add to users' experience simply by incorporating a new and pleasurable sensory dimension.

Implementation

The sound system is directly triggered by the emotional model. One set of sound samples illustrates the emotional levels of the 4 strongest emotions. The other set of sound samples give direct feedback when there has been a significant change in emotional levels, i.e. a postcard has been input by the user.

The sound system is implemented with a set of rules that map emotions and changes in emotions to particular sound files. The sound system regularly checks the emotional levels in the emotional model to see if any of the pre-configured conditions for a sound response are met. If so, it executes the appropriate sound control commands.

The sound player contains dedicated channels for each pair of opposite emotions, e.g. anger/peace. Every channel only plays the audio representations related to these emotions. The number of emotional audio representations playing at the same time is restricted to a configurable value (default is 4).

SOUND DESIGN

The choice of audio material for expressing emotion is based on the need for the system to express both the overall emotional state of the Influencing Machine as well as significant changes in that state. We decided to use mainly configurations and citations from modern, nearly "non-tonal" concert music, some of which we have freely edited or adapted. By doing so, we are able to draw on a long and continuing tradition of experience relating affects to audio structures, a tradition stretching from baroque music until today.

One of the most attractive characteristics of this tradition is that many of its structures permit a free relationship in coordinating the replay of samples without losing affective impact, a feature that supports real-time selection and coordination of sound clips. Another aspect that is fortunate for us is that this kind of free music since Pierre Boulez uses the articulation and body sound of the instrumentalists as part of the music. This audio material therefore contains a lot of expressive body sounds, especially different kinds of breath with evocative characteristics. This material was supplemented by recordings of natural sounds and human voices in different environments.

The sounds that express emotions differ in regard to their origin (natural sounds, different traditional instruments and electronic sound) and to their composition. In most cases, these sound representations differentiate not only the type but also the intensity of emotions by, for example, changing volume, using familiar melodies or by other strategies of audio variation (e.g. variegated rhythm, melody or harmonies, set filters, modify formants). The overall impression of the sounds expressing emotional state recalls strategies of mixing up sound in early tape music, e.g. in the school of Pierre Schaeffer or especially in the few examples of surrealistic music, where tape recordings had been synthesized with digital sounds and composited instrumental parts [20].

THE SIGNALS

The "signals" illustrating significant changes in emotional value, whose main purpose is to provide immediate feedback for the input system, are short, typical sounds or phrases. Their sound design is motivated by the need for the kind of sounds to be related to (or at least musically sensible with) the established audio expressions for the state of the system. They differ in instrumentation, in volume, timbre, rhythm and harmony and they can differentiate the direction of change through melodic elements and articulation. For example, changes in cold are expressed through a two-tone melody or interval that goes up with increases in cold and down with decreases. Similarly, changes in hot are expressed by a glissando going up as the value increases and down as the value decreases.

In testing the sound system we found that the signals do not sound simply like disparate sequences or superimpositions, but also display synergy. That is, they seem to "fuse" with each other and the emotion sounds, if the several emotion values simultaneously change in a significant way, e.g. if the value of them 'jumps'. So the result of audio output is that the signals show the tendency to built common phrases, providing the impression of an organic process. In addition, though the sounds were developed largely independently of the graphical output, sound and graphics combine well, creating a feeling of synergy in the behaviour of the system.

EVALUATION

Pia Mårtensson and Kristina Höök conducted an initial formal evaluation of the Influencing Machine without the sound system. The results of this evaluation were used to inform subsequent design of the Influencing Machine system.

Setting of the Influencing Machine at SICS

In the initial evaluation, users were brought in small groups (six groups with in total 12 subjects) into a room with the Influencing Machine. Users were told that the installation had something to do with emotions, and were then allowed to play with the system as long as they liked. On average, they spent about 20 minutes in the room.

Generally speaking, users were first curious, then became frustrated. Often this frustration stemmed from not being able to control the machine. They had a great deal of trouble figuring out the relationship between postcard and drawings. For some users this became a barrier that stopped their interest in the machine. For others, it was a challenge:

> "It changed a bit over time. At first I felt the machine influenced me before I realized I might be able to change it the other way around. The machine drew a lot of gray, then I would put in cards that had gray in them. If it drew parallel lines I picked industrial motives, e.g. I tried to give it what I thought it wanted. After a while I realized I could think of it as a child that you can coach and that's when you can change it the other way around. It drew a gray circle… ok! But have you thought of that you could draw little red spots as well? Communicate something.. like… have a look at this card! In that way you can encourage maybe.. you put in cards and it draws something new. A little anyway, it felt like it was the machine that was drawing but it got some inspiration from the cards."

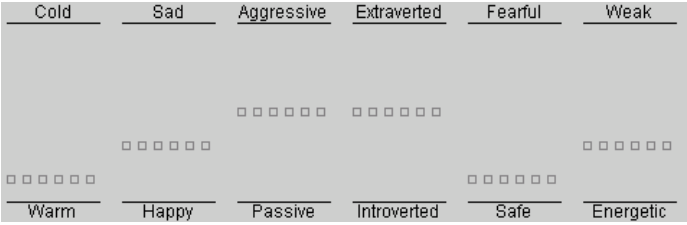
Some users found the Influencing Machine drawings too simple and drawn too slow. One user says: "[I found it] somewhat irritating as it painted so slow." Others found this to be like "watching clouds", thus a soothing, relaxing experience. In general, subjects did not understand that the drawings symbolised feelings, while they were more inclined to interpret the postcards in terms of emotions: "The ones that portrayed something were the best. Can communicate emotions much better. The one-colour I could not communicate as much with." The mailbox itself was liked. One subject said: "It was so much fun, you got the urge to put down your fingers and see if there was some kind of animal in that box." Unfortunately, the bar code reader made a beep whenever a postcard was inserted. This led subjects to think of the mailbox as a machine rather than a form of communication with a semi-living being.

One of the goals of the Influencing Machine is to incur critical reflection in users. This was sometimes successful, sometimes not. "I do not know, but I shall not be able to sleep tonight I think, because I shall be thinking about WHAT the whole point was really? I do not know why really... This issue with the possibility to influence it and that. One will think about this for a long time." Another user: "Well...., it does raise questions or a 'wonder' on what it was supposed to mean. You look for a purpose". Another user responds: "I shall sleep well!"

It is difficult to know if this is a success or failure. The concept of influencing instead of controlling a machine may be a difficult one to get; perhaps one cannot expect many users to be able to understand it. An additional complication is the frustration that users often developed with lack of control. Many users got irritated and frustrated when they could not figure it out. Certainly this is an affective reaction, but not one that we had intended. On the positive side, this may spur them to develop a different perspective on their interaction with the system. On the negative side, we had not intended to build an intellectual torture device. These thoughts and observations led to the following system design changes.

Design Changes from Evaluation

Users were confused about the emotional meaning of the imagery. The addition of the sound system helps to clarify the agent's interpretation of input cards and its emotional state. In addition, we developed an internal emotional display, which shows the level of each of the internal emotions. Although we were reluctant to show these internals, by offering the user an opportunity to understand how the agent is designed to feel, users can and do engage in critical reflection on whether they believe that the drawings actually express the stated internal emotion state. Ideally, this display will fade away over time, supporting users through their initial exploration without constraining further interaction. The system can be shown either with or without this display, allowing for further experimentation.

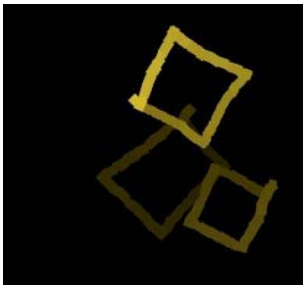


Part of the emotion display.

Users were also confused about the nature of influencing versus controlling the system. With the above improvements to emotional expression, including direct sound feedback instead of mechanical Mailbox beeping for changes in emotion, users will hopefully have a better understanding of how they affect the system. At the same time, this concept is subtle and runs counter to users' everyday experiences with computers; it may simply be in its nature that it is hard for users to understand.

In part because of these problems, users seemed unsure of the point of the system. In the next evaluation, we will also provide users with more background information, including a brief statement on affective computation and the suggestion that they ask themselves which emotions they can see in the system.

Finally, users were sometimes bored by the drawings themselves. Speeding up the drawings, reducing the persistence of behaviours so that new forms appear more quickly, and adding some more complex drawings will probably raise user interest. Also, transitions between drawings need to be handled more gracefully. In the old version, the system draws for a while and then clears the screen and starts over. We have now reimplemented the graphics to remove these rough breaks by layering over one another and gradually fading away.



Images appear and gradually fade away, becoming overlaid with newer scribblings.

Informal user tests and public demonstrations incorporating some of these design changes suggest that they have substantially improved the comprehensibility and enjoyability of the system. The sound system in particular has added a new dimension to the experience. A final formal evaluation of the improved Influencing Machine by Gerd Andersson and Kristina Höök is scheduled to begin shortly.

DISCUSSION

Initially, we conceptualised this project to be mostly about exploring the agent-human relationships, in particular the question of influencing vs. controlling an artificial, emotional agent. We presumed the affective parts would be relatively straightforward, an implementation detail. Instead, we ended up wrestling with the enigmatic and elusive nature of affect, struggling to avoid either rationalizing and stultifying it, or, on the other hand, having the whole interaction disintegrate into incomprehension. The Influencing Machine ended up influencing us: we gradually learned to appreciate and harness the enigmatic nature of affect. In the process, we learned (and relearned) the following lessons about the enigmatics of affect.

Emotion in interaction refers not only to whether the system is able somehow to apprehend the user's emotions, but also to the emotions that the user develops in interaction with the system. Interfaces may be built to engage user's emotions not only by being an input device for emotions, but by setting up an interaction situation that leads the user to feel emotionally engaged. An important aspect of such interaction is tangibility. Physical interaction is, often, emotional interaction, and being able to feel pieces of the system can easily give rise to emotions in the user. We do not believe our system would work as well if postcards appeared on a screen and were clicked on. The physicality of the system brings it into the user's world and invites everyday engagement.

More generally, the design of systems that support enigmatic affect must include considerations of aesthetics, poetics, and meaning. Human, cultural aspects, which are often difficult to keep into account in a purely technical approach, are essential for achieving a truly emotional interaction. For example, the look-and-feel of an application, and not just its underlying algorithms, will have a large

impact on emotions users will feel while engaged with the system. This includes consideration of the metaphors used in interaction. For example, Chris Dodge's interactive networked installation The Bed consciously draws on emotionally evocative metaphors of breath, heartbeat, and movement to create an intimate, shared, and emotionally meaningful space [13]. As systems like this make clear, emotional interaction can be achieved by selecting emotionally-laden metaphors of interaction. This is a challenge for a computer science that is opening itself up to affect, since it has traditionally been reluctant to draw on humanist perspectives for its work.

Most fundamentally, the enigmatics of affect taught us that the emotion is in the interaction, not (necessarily) in the code or hardware. Emotional interaction design includes the real (not represented) emotions of the user, and may call them forth without explicitly - or even correctly - modelling them.

These principles and the technologies they led to are closely related to research in believable agents, which seeks to understand how to build personality-rich computer characters which are not necessarily intelligent but can be narrated as such. We were particularly inspired by Bates, Loyall, and Neal Reilly's work on discretised emotion as a placeholder token for generating behaviour that appears emotional [3] [21] [26]. We build on research exploring ways to construct agent behaviour that explicitly consider and support comprehensibility [34] [32], but do not allow for flexible interpretability. The Influencing Machine moves in the direction of Simon Penny's Petit Mal [29], a robot built to explore the interpretability of artificial agents. Petit Mal's body is structured so that, though the agent's behaviour is simple, the dynamics of its movement will be complex.[5] The complexity of this movement invites complex interpretation of its activity; users regularly narrate the robot's activity in ways that far exceed the robot's 'actual' (internal) functioning.

With such interpretational openness we move towards something we might call enigmatic computing, or the paradoxical creation of well-defined internal structures that support open-ended, unstructured experiences of the system. Michael Mateas has devised ways to conceptualise such design through the notion of authorial and intepretational affordances [23]. Interpretational affordances are the natural opportunities for interpretation the system provides the user, often from the domain of art, design, and the humanities. Authorial affordances are the natural opportunities the system provides authors for expressing their intentions through the system, often from the domain of technical practices. Artefacts that are intended to combine complex technology (authorial affordances) with meaning-making (intentional affordances) must be constructed with both these aspects continually in mind. In the Influencing Machine design process, this was done by moving continually between determining the internal model and technologies of emotion and the opportunities these technologies gave users for open interpretation.

[5] The need for a poetics of design is reflected in Penny's work by Petit Mal's skin of flocked, flowered tablecloth, which turns an industrial-looking robot into a gangly, lovable geek.

CONCLUSION

What is actually influenced in the Influencing Machine? Just as users interpret the images generated by the machine, input also requires an act of interpretation. The Mailbox not only influences the emotions of the agent, it also influences the perception of the user and creates an open space of interpretation. The Influencing Machine is an installation and an experiment. It is an experiment for us as researchers, because we can invite users in to see how they relate to the system and think of it. It is also an experiment for users, because it offers them the opportunity to try out and play with the notion of affective computing, developing their own perspective on it. Finally, it is a cultural probe [16] in the field of affective computing itself, a test case for a new approach to affect that incorporates humanist concepts of emotion and privileges the lived experience of emotion over its informatics.

CONTACT

The Influencing Machine was developed at the MARS Exploratory Media Lab, Fraunhofer Institut für Medienkommunikation, Schloss Birlinghoven, 53754 Sankt Augustin, Germany. MARS is led by Monika Fleischmann and Wolfgang Strauss. Except as noted, MARS authors can be reached at firstname.lastname@imk.fhg.de. Evaluation was done at the Human Computer Interaction and Language Engineering Laboratory, Swedish Institute of Computer Science, Box 1263, SE-16429 Kista, Sweden. Pia Mårtensson is at pima@sics.se; Kristina Höök is at kia@sics.se.

ACKNOWLEDGMENTS

Thanks to Gabriele Blome, Ulrike Boecking, Monika Fleischmann, and Wolfgang Strauss for their help in the development of the project. Florian Dombois, Thorsten Holtkämpfer, and Leonie Schäfer graciously played guinea pig for informal user testing at GMD.

The Influencing Machine project is funded by the EU IST program as part of the SAFIRA Project on affect in interaction, which is led by Ana Paiva. Many thanks to our project partners for their help, input, and support.

REFERENCES

1. Agre, P. *Computation and Human Experience*. Cambridge UP, Cambridge, 1997.
2. Ark, W., Dryer, D., and Lu, D. The emotion mouse. *Proceedings of HCI International 1999*. Munich, Germany (August 1999).
3. Bates, J. The role of emotion in believable agents. *Commun. ACM*, Special issue on Agents, 37, 7 (July 1994), 122-125.
4. Bettelheim, B. *The Empty Fortress: Infantile Autism and the Birth of the Self*. Free Press, New York, 1972.
5. Blumberg, B. Action-selection in Hamsterdam: Lessons from ethology, in *Proceedings of the 3rd International Conference on the Simulation of Adaptive Behavior* (Brighton, 1994).
6. Blumberg, B. *Old Tricks, New Dogs: Ethology and Interactive Creatures*. PhD thesis, MIT Media Lab, 1996.

7. Brooks, R. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2, 7 (April 1986),14-23.

8. Brooks, R. A. Intelligence without reason, in *Proceedings of IJCAI-91 (*Sydney, Australia, August 1991), 569--595.

9. Brooks, R. Integrated systems based on behaviors, in *Proceedings of AAAI Spring Symposium on Integrated Intelligent Architectures*, Stanford University (March 1991). Also available in *SIGART Bulletin*, 2, 4 (August 1991).

10. Burton, E. Artificial innocence: Interactions between the study of children's drawing and Artificial Intelligence. *Leonardo*. 30, 4 (1997), 301-309.

11. Burton, E. Representing representation: Artificial Intelligence and drawing. *Computers & Art.* Intellect Ltd, . Exeter, 1997.

12. Cohen, H. The further exploits of AARON, painter. Stanford Electronic Humanities Review, Special Issue on Constructions of the Mind, 4, 2 (1995).

13. Damasio, A. *Descartes' Error: Emotion, Reason, and the Human Brain*. Gosset Putnam, New York, 1994.

14. Dodge, C. The Bed: A Medium for Intimate Communication., in *Proceedings of CHI '97* (Short papers). http://www.acm.org/sigchi/chi97/proceedings/short-talk/cd.htm.

15. Fernandez, R., Scheirer, J. and Picard, R. Expression glasses: a wearable device for facial expression recognition. MIT Media Lab Tech. Rep. 484, 1999.

16. Gaver, B, Dunne, T., and Pacenti, E. Cultural probes, *interactions* 6, 1 (Jan-Feb 1999), 21-29.

17. Heller, E. *Wie Farben Wirken: Farbpsychologie, Farbsymbolik, Kreative Farbgestaltung,* Rowohlt Verlag, Reibek bei Hamburg, 1989.

18. Hoffman, E.T.A. Der Sandmann. *Der Sandmann / Das Öde Haus*. Reclam, Stuttgart, 1969.

19. Kellogg, R. *Analyzing Children's Art*. Mayfield Publishing Company, Mountain View, CA, 1970.

20. Liesendahl, R. Andenken an Einzelnes, *José-Luis de Delás, Musikkonzepte 78*, ed. H.-K. Metzger, Munich, Germany, 1992.

21. Loyall, A. B. *Believable Agents: Building Interactive Personalities*. PhD Thesis, Carnegie Mellon University Dept. of Computer Science, 1997, CMU-CS-97-123.

22. Maes, P. The dynamics of action-selection. In Proceedings of IJCAI-89. Detroit, MI: Morgan Kauffman (August 1989), vol. 2, 991-997.

23. Maes, P. How to do the right thing. *Connection Science Journal*, 1, 3 (1989), 291-323, 1989.

24. Mateas, M. Expressive AI. SIGGraph 2000 Electronic Arts and Animation Catalog (New Orleans).

25. McCorduck, P. *AARON's Code: Meta-Art, Artificial Intelligence, and the Work of Harold Cohen*, W.H. Freeman and Company, New York, 1991.

26. Neal Reilly, W. S. *Believable Social and Emotional Agents*. PhD Thesis, Carnegie Mellon University Dept. of Computer Science, 1996, CMU-CS-96-138.

27. Ortony, A., Clore, G.L. and Collins, A. *The cognitive structure of emotions*, Cambridge University Press, N.Y, 1988.

28. Pearson, D., Huffman, S., Willis, M., Laird, J., and Jones, R. Intelligent multi-level control in a highly reactive domain, in *Proceedings of the International Conference on Intelligent Autonomous Systems* (Pittsburgh, PA, 1993).

29. Penny, S. Agents as artwork and agent design as artistic practice, in *Human Cognition and Social Agent Technology*, ed. Kerstin Dautenhahn, John Benjamins Publishing Company, Amsterdam, 2000.

30. Picard, R. *Affective Computing*, MIT Press, Cambridge, 1997.

31. Richter, H,.-G. *Die Kinderzeichnung: Entwicklung, Interpretation, Ästhetik.* Cornelsen, Berlin 1997.

32. Scheirer, J., and Picard, R. Affective Objects. MIT Media Lab Technical report, nº524.

33. Sengers, P. Designing comprehensible agents, in *Proceedings of IJCAI-99* (Stockholm, Sweden). Morgan Kaufmann, Vol. 2, 1227-1232.

34. Stone, B. A. Dynamically sequencing an animated pedagogical agent, Proceedings of AAAI-96, Portland, OR, 424-431.

35. Strauss, M. *Von der Zeichensprache des kleinen Kindes: Spuren der Menschwerdung*. Verlag Freies Geistesleben, Stuttgart, 1994.

36. Tausk, V. The influencing machine, in *Incorporations*, ed. J. Crary and S. Kwinter. Zone Books, New York 1992.