

# Machines With a Different Calling

Marc Böhlen <sup>1</sup>, Michael Mateas <sup>2</sup>

<sup>1</sup>*Department of Media Study, SUNY Buffalo, USA, mrbohlen@acsu.buffalo.edu*

<sup>2</sup>*School of Computer Science, Carnegie Mellon University, USA, michaelm@cs.cmu.edu*

## Abstract

*This paper presents an argument for including concepts from the social sciences and the arts in the design of robots for intimate social environments. An example of a robot designed as an interactive dining table situated in a restaurant is described and certain aspects of the particular design approach are explained. It is shown how this can help think about ways to integrate robots into cultural spaces.*

## 1. Introduction

If robots are to become ubiquitous in personal spaces they must be able to deal with complexities that arise in real social spaces. If robots are to become ubiquitous, accepted and desirable as companions they should become sensitive to all facets of human life. In order to attempt to achieve such a difficult task, an interdisciplinary approach that includes knowledge from the social sciences as well as the technical sciences should be employed. Ethnography and sociology have as much to contribute to this agenda as AI and machine learning. Furthermore, activities in the robotic arts are suggested as inspiration nodes for conceiving robots for complex social situations. Similar thoughts have been expressed in the AI community where some computer scientists have established a research practice [16] that includes what Philip Agre [1] calls a critical technical practice. This approach links development within computer science to the expression of aspects of human experience by including ideas common in cultural theory into the AI research agenda. The Human Computer Interaction community has readily embraced perspectives from the social sciences and the arts [5]. Additionally, knowledge gained in ethnographic studies has been included in the evaluation of

computing needs [10]. In short, there is a clear tendency in these fields to include cultural frames of analysis when addressing issues concerning human beings and technological devices.

## 2. Related Work

While the robotics community has to date paid little attention to the cultural context of its production, the idea of a social role for robots is well established. There has also been research in making intuitive interfaces to facilitate interactions between robots and humans. A number of museum tour-guide robots by Nourbakhsh [12] and Thrun [18] have demonstrated this with impressive results. Caretaker robots [15] have been designed to assist and interact in intuitive ways with technically non-inclined elderly people. In addition to socially acceptable interfaces for robots there is also research on socially acceptable behavior for robots. Simmons [17] shows how a robot can identify a line of waiting people and queue up. Dautenhahn [7] has advocated the use of heuristics from the human sciences in the design of autonomous agents and machine vision systems. Picard [13] and the Affective Computing Lab have laid the foundation for integrating affect and detectable emotions into computational processes.

## 3. When artists make robots

There is an interesting history of integrating machinery into artwork, too long to properly unfold in this short paper. Marinetti [8] and the Futurists in Italy in and after 1909 were probably the first to embrace industrial machinery. For the Futurists, industrial machinery embodied a new aesthetic of speed and social change. Artists' attraction to machinery and mechanization continues today.

There is an active community of artists who employ robotic technologies into their work.

Unfortunately, the relationship between artists and roboticists is a strained one. Roboticists generally find artists' robotic work to be amateurish, technologically simplistic and uninformed by current state of research. In the context of this paper, artists' approach to robotics is of interest because their practice is inherently informed by and dependent on knowledge from the human sciences. Artists generally contextualize their work in a large cultural context and tend to be fluent in cultural theories. When these sensitivities are carried over to robotics related work, different types of questions concerning the role of robotics are formulated. Here are two examples.

The Institute for Applied Autonomy [2] has created a series of robots that address ideas surrounding the impact of automation and robotics on freedom of speech. Their best-known work, *Graffiti Robot*, is a remote controlled mobile robot that can be programmed to spray short texts onto sidewalks and streets. While the means employed to create this machine are very simple, the question that the work raises is more subtle. If a robot sprays graffiti onto a street, who is responsible for its content?



Figure 1: Norman White, "Helpless Robot, 1996"

Norman White [11] is another artist-roboticist who has done very interesting work over the years. His *Helpless Robot* encourages peoples to move it, as it would "like". The machine attempts to assess and predict human behavior such that a person can help it. This robot is of interest for the particular approach to robot-human exchange it suggests. Can

we build machines that politely anticipate our behavior? What are the best exit strategies, on an interaction level with human beings, when robots malfunction? Can malfunctioning robots generate empathy for themselves? Research robots avoid these questions; when research robots fail, they are quietly ferried back to the workbench.

#### 4. Contemplative Robotics

The growing interest within the robotics community in embedding robots into all levels of society is the rational for including thought processes from the arts into mainstream robotics research. The marriage between the two practices, robotic research and art practice, can build a basis for a critical technical practice that combines intuition with technical and cultural understanding. This combined practice of robot design is called *Contemplative Robotics* (CR). It attempts to engage people, as good art has the power to do, and address in cultural contexts exchanges between humans and robots. Typically this happens in complex and sometimes delicate social environments in the unstructured real world. As opposed to a pure art practice, CR additionally attempts to find viable technical solutions to the problems it investigates. The following section describes a test bed for some of these ideas in a current research project.

##### 4.1 Case study: The DinnerTable Project

**Overview.** The DinnerTable Project is an ongoing effort that comprises activities and research in biometric data acquisition, and classification, kinematic sculpture, automated narration and CR.

DinnerTable (DT) is an interactive robotic narration machine, in the guise of a generic table that observes conversation and activity around the table and generates narratives in the form of an interactive game in response to the observed activity. The table is situated in a restaurant where people meet for a fine dinner. Screens integrated into the tabletop depict both a menu and instructions for use when guests sit by the table. If the guests choose to engage with the table as a robotic companion, they can do so by picking one of four game pieces and putting it on the table. Each piece represents a character with preexisting dispositions and a background of recorded stories that condition the game to be played. In response to the conversational flow throughout dinner, the pieces move on the glass tabletop, designed as a game board of an urban

environment, and enact a unique story that is jointly created by the users and the table in real time. The story is conveyed through a mixture of game piece movement, projected imagery and text. The users are players and authors at once as the table creates narratives informed by the dynamics of speech and the activities of the guests at the table.

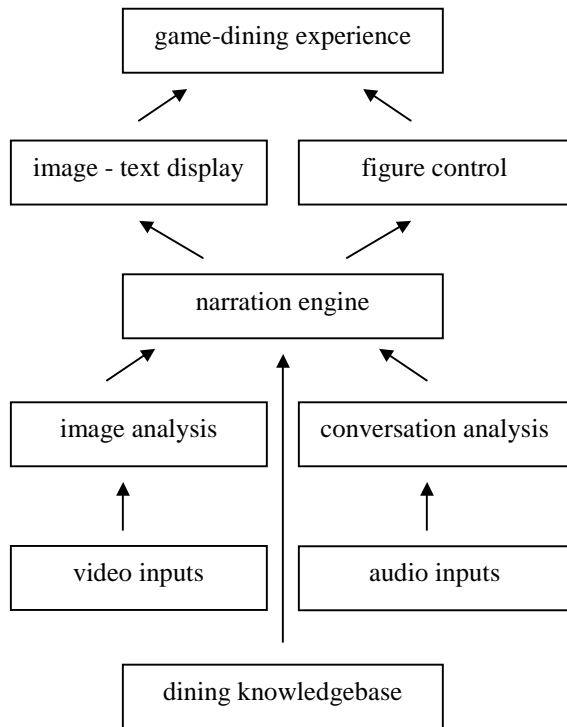


Figure 2: Schematic overview



Figure 3: First prototype of DT

**Figure control.** The manipulator must be able to move two game pieces simultaneously through the urban game board landscape. We can achieve this with two planar tracks that are in turn mounted in a

larger track. Each of the smaller tracks has a left or right field of operation that overlaps with that of the other track. By this design, a game piece traveling from one edge of the board to the other is handed off from one track to the other, invisible to the table guests. The motions of the manipulator are transferred to the game board pieces via electromagnets. A rigid aluminum frame provides the required stability.

**Image analysis.** The image analysis module makes use of the matrox MIL imaging library for image acquisition and low level processing. Higher-level tracking algorithms in C++ are custom built from the library's primitives. A single CCD video camera is mounted on the ceiling. Activities around the table, such as the appearance or disappearance of dinner guests and restaurant staff, can be determined by standard image analysis techniques of arithmetic operations, filtering and segmentation. This is possible because we make use of the ritualized dining-context for useful prior knowledge (such as the fact that a waiter comes to the table before the meals are served) to reason about ambiguous results. The serving of a glass of wine, a toast and a kiss between the guests can be detected. Furthermore, an additional tracking module periodically checks the positions of the game board pieces and compares results with the driver electronics responsible for controlling the 2D manipulator. There is currently no gesture recognition in the system but a second side view camera dedicated to this task is being tested.

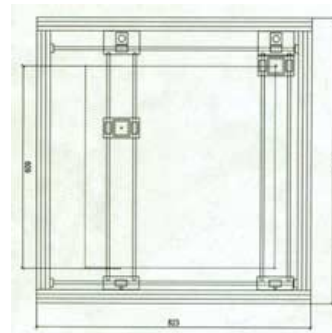


Figure 4: Dual planar manipulator for figure control

**Conversation analysis.** The conversation analysis module is written in Matlab using the Data Acquisition Toolbox that allows simultaneous sampling of analog inputs on multiple channels. This project requires a method of tracking the discussion dynamics during the dining event.

Initially the goal included a semantic relationship to the content of the visitors' discussion. This proved a bad idea. First, not all table talk is nutrition for an interesting story. Second, very large vocabulary, speaker independent, multi-language, continuous multi-channel real-time conversational speech recognition in noisy environments, while highly desirable, is still beyond the state of current speech recognition research. Current speech recognition technology can handle natural speech, but only in a limited domain or with a high word error rate. Making false links between table discussion and generated narratives would be alienating to guests. Hence we have decided to remove the link to the semantic level of the conversation and concentrate on the mood and timing of speech dynamics. As a consequence, the narration engine receives a higher level of autonomy (see below). Since it is the flow of conversation dynamics we seek, we choose to analyze spoken language on a signal processing level. This approach can nonetheless deliver acoustic characteristics of speech [14] that can then be used in conjunction with input from the vision system to track the high-level interaction dynamics of interest. The signal processing analysis approach allows quick analysis on two inputs in multiple languages and can include infant and pet directed speech [6]. For this project we are interested in keeping the rich input data "intact" and remaining sensitive to the context in which it is generated.



Figure 4: Second prototype of DT (CAD)

The current approach allows us to detect, with two noise reducing microphone arrays, laughter, change of tone, coughing, silence, and excitement via frequency and amplitude analysis of the speech signals over time. The frequency domain delivers information on the change of tone and on the vocal tract resonant frequencies, formants 1 and 2. Normalized power densities deliver intensity and excitement. Not all the above parameters are equally robustly detected. We are in the process of evaluating additional strategies of improving the system's performance and including additional salient parameters such as whispering. The classification of the signal level information is currently performed via a back propagation neural net. A first round of training data has been collected from multiple speakers in two languages (English, German). The outputs from the classification algorithm, collected simultaneously from two users, open a window into dinner discussion dynamics.

**The narration engine.** We are investigating the aesthetics of *assisted making* of interactive poetic narration. Our machine questions the role of human authorship and ponders the potential of internal computational procedures for the capacity of narrative significance. We have "delegated" top-level authorship to a selection of 6 seminal narration strands from literature and film and have created a database of over 2000 images to represent them. From this database DT creates and recombines particular narration instantiations as desired by its internal constraints and external stimuli described above. The AI system that is responsible for this, the narrative engine, must satisfy the following constraints:

1. Be clearly responsive to the social situation.
2. The participants should be able to discern the goal that the game pieces have and the rules within which the pieces attempt to achieve this goal.
3. The sequence of events produced by the table should form a narrative; the current event must be constrained by the history of past events and move towards a climactic future event to convey theme.

This engine must provide the ability to map from sensory inputs to triggered sequences of game piece movements, images and text, provide support for these sequences to interact, such that the sequences can be perceived as moves in a game, and organize these mappings into a script-like global structure which constrains the mappings such that interactions

with DT form narratives. The solution adopted by DT is to use reactive planning, specifically the reactive planning language ABL, which is based on the reactive-planning language hap [3, 9]. An ABL program is organized as a collection of independent behaviors, each of which consists of some number of primitive steps, which move pieces and display images on the table, computational steps which perform arithmetic and logical computations, and subgoal steps, which attempt to accomplish a more complex goal by appropriately selecting other behaviors to work on the goal. Additionally, an ABL program has a working memory where it stores working memory elements (WMEs). These working memory elements can contain whatever information the program needs to keep track of. Behaviors and goals can test working memory to determine which behaviors are selected, and whether goals and behaviors succeed and fail.

WMEs can be tied to sensory input. In this way, an ABL program can use WMEs as a window onto the world. For example, a WME could keep track of how much laughter (on some numerical scale) has occurred at the table in the last 30 seconds. The WME is maintained by a process external to the ABL program, which senses laughter at the table. This WME can then be used to influence behaviors. ABL fulfills the narrative engine's requirements listed above in the following way:

1. The hierarchical structure of behaviors provides the capability to control the global structure of events, through top-level sequential and parallel behaviors, so as to make a narrative happen.
2. The dependence of behavior execution on tests of sensed WMEs provides the ability to allow details of the narrative to depend on the sensed situation at multiple levels of abstraction so as to provide a non-human commentary on the social situation.
3. The ability to maintain long term persistent goals which are pursued reactively and opportunistically supports structuring reactivity in terms of helping or hindering character goals so as to create a game-like situation.

**Non-intrusive interaction.** DT is an exploration into non-intrusive methods in recording and evaluating biometric data. Both the choice of our biometric parameters as well as the handling of the processed results writes this into the project's agenda. There are no buttons to touch, no questions to answer, and there are no wires connected to the guests. The input parameters are image and sound

based. Acquired data is fed into the machine in a closed loop manner and becomes nutrition for the narration engine.

**A situated robot.** DT lives in the intimate setting and intimate exchanges that can occur in dining situations. The work would like to add to the dining dynamics and humbly feed off it while narratively contributing to the private exchange. This requires a delicate form of responsibility. It is essential in CR to address and include cultural parameters into the conception of machinery designed to interface with human beings. It is important to be able to react in one fashion in a boisterous conversation during a dining situation involving multiple courses and in a different fashion in a quick discussion during a snack at a teahouse.

**A robot you can trust.** Dinner guests have the option of disengaging the table's narrative companionship. If they desire to be alone and undisturbed, DT will revert to its unanimated state, that of a generic table. This *duality* is an attempt to leave the visitors in complete control and foster a notion of trust between the dinner guests and the robot. The robot can disappear into the background [19] and the dinner event can happen as if the robot did not exist. However, if guests feel comfortable enough to test their curiosity and engage DT's companionship, they can rest assured that no data will be retained or stored in the system. All collected data, audio as well as visual, is fed directly back into the narration generation in a closed loop form. People are very sensitive to surveillance and privacy issues. DT instantiates its own ethic in this regard by employing *soft surveillance* that prevents data misuse on the design level. DT gives back what it takes. These features are understood not as auxiliary appendages but essentials that inform this robot design from beginning to end.

## 5. Future work

We are in the process of designing a second DT prototype that will allow us to move our tests from the laboratory to a real restaurant. As mentioned above, the addition of a second, side view camera dedicated to gesture recognition is being evaluated. Improvements in the audio signal analysis system and development in the narration engine are ongoing. DT, as an instantiation of CR, is also becoming a conceptual catalyst for thinking differently about robots.

## 6. Thinking differently about robots

CR attempts to help find solutions to the integration of robotic technology into human culture spaces by including intuition from the arts and knowledge from the social sciences into its design parameters. Like tourists in a foreign country, the most sophisticated robots we can build will need guidance in intangible aspects of life in our complex multiethnic and pluralistic societies. The richness in current technological innovation should be coupled explicitly with awareness of the cultural contexts in which it operates. We can do more than build efficient machines. We can create robots that are acceptive of human follies and fears, robotic systems that understand that the human notion of time and duration [4] is different from clocked machinic time. Expanding the robotics agenda in such ways can help develop better robots to live with in the long term.

## 7. Acknowledgements

The DinnerTable project is funded in part by a grant from the Institute for Studies in the Arts (ISA) in Arizona.

## 8. References

- [1] Agre, P., "Toward a critical technical practice: Lessons learned in trying to reform AI". In: Geoffrey C. Bowker, Susan Leigh Star, William Turner, and Les Gasser, eds, *Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide*, Erlbaum, 1997.
- [2] <http://www.appliedautonomy.com/>
- [3] Bates, J., A. B. Loyal, A. B., Reilly, W. S., "An Architecture for Action, Emotion, and Social Behavior". Tech. Report CMU-CS-92-144, Carnegie Mellon University, 1992.
- [4] Bergson, H. *Creative Evolution [L'Évolution créatrice, 1907]* Trans. from the French by Arthur Mitchell, 1923.
- [5] Böhlen, M., "A Different Kind of Information Appliance: Fridge Companion". Proceedings of ACM CHI 2002 Conference on Human Factors in Computing Systems, Minneapolis, 2002.
- [6] Burnham, D., Francis, E., Vollmer-Conna, U., Kitamura, C., Averkiou, V., Olley, A., Nguyen, M., & Paterson, C. "Are you my little pussy-cat? Acoustic, phonetic and affective qualities of infant- and pet-directed speech". In: R. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5<sup>th</sup> International Conference on Spoken Language Processing: ICSLP'98, 1998, Volume 2*, p. 453-456.
- [7] Dautenhahn, K., Ogden B., "Embedding Robotic Agents in the Social Environment". The 3rd British Conference on Autonomous Mobile Robotics and Autonomous Systems, Manchester, 5th April 2001.
- [8] Marinetti, F., T., "Manifesto del Futurismo", *Le Figaro*, Milan, 1909.
- [9] Mateas, M. and Stern, A., "A Behavior Language for Story-Based Believable Agents". In: *Working Notes of Artificial Intelligence and Interactive Entertainment*, Eds, Forbus, K and Seit, E., AAAI Spring Symposium Series, Menlo Park, 2002.
- [10] Mateas, M., Salvador, T., Scholtz, J., and Sorensen, D., "Engineering ethnography in the home". In: *Proceedings of CHI'96*, ACM Press, 1996, p. 283-284.
- [11] <http://www.normill.com/ntwcv.html>
- [12] Nourbakhsh, I., Willeke, T, Kunz, C., "The history of the Mobot museum robot series: An Evolutionary Study". In: *Proceedings for FLAIRS*, 2001.
- [13] Picard, R., "Affective Computing", MIT Press, Cambridge, 1997.
- [14] Pereira, C., Watson, C., "Some Acoustic Characteristics of Emotion". In: *The 5th International Conference on Spoken Language Processing*, Sydney, 1998, p. 927- 930.
- [15] Roy, N., Thrun S., et al., "Towards Personal Service Robots for the Elderly". In: *Proceedings of the Workshop on Interactive Robots and Entertainment (WIRE)*, Pittsburgh, 2000.
- [16] Sengers, P. "Practices for Machine Culture: A Case Study of Integrating Artificial Intelligence and Cultural Theory." *Surfaces*. Volume VIII, 1999.
- [17] Simmons, R. Nakauchi, Y., "A Social Robot that Stands in Line". *International Conference on Intelligent Robots and Systems, IROS*, 2000.
- [18] Thrun S., et al., "Experiences with an interactive museum tour-guide robot". In: *Artificial Intelligence*, 114 (1-2): 3-55, 1999.
- [19] Weiser, M., "The Computer for the 21<sup>st</sup> Century". In: *Scientific American*, 265, 3, 1991, p. 66-75.