Embodied Cultural Agents: at the intersection of Robotics, Cognitive Science and Interactive Art

Simon Penny
Associate Professor of Art and Robotics, Carnegie Mellon
CFA 300, 5000 Forbes Ave Pittsburgh PA 15213 USA
penny+@cmu.edu

Abstract

This paper outlines the development over several years of Petit Mal, an autonomous robotic artwork, and discusses a new project arising from it. Central concerns are an holistic approach to the hardware/software duality, the construction of a seemingly sentient and social machine from minimal components, the generation of an agent interface utilising purely kinesthetic or somatosensory modes which 'speak the language of the body' and bypasses textual,verbal or iconic signs. General goals are exploration of the 'aesthetics of behavior', of the cultural dimensions of autonomous agents and of emergent sociality amongst agents, virtual and embodied. The research emerges from artistic practice and is therefore concerned with subtle and evocative modes of communication rather than pragmatic goal based functions. A notion of an ongoing conversation between system and user is desired over a (pavlovian) stimulus and response model. The paper concludes with a description of the project Caucus, a group of Petit Mal-style robots which will generate sociality on-the-fly as a result of the exhange of linguistic tokens.

I began to design "Petit Mal: an autonomous robotic artwork" in 1989. I must emphasise that at the time, my familarity with robotics was minimal, I was propelled to this task by a desire to build a more sophisticated interactive artwork. The project arises out of an artistic context and work proceeded in the style of an artistic project: it was radically underfunded; it was based on the presumption that if I could imagine the end product, I should be able to make it. By 1992 I had mustered enough funds to begin the project. Over about 3 years I built it, essentially single handedly, with the help of several sucessive undergraduate assistants, for around $5000. Three years later, in february 1995, Petit Mal made its public debut. Since then it has proven to be reliable and robust, it has been shown in many festivals where it must interact with the public continuously for 8 hour days, for weeks at a time.

At the outset, I did not describe the project as an embodied agent, I was unfamiliar with that terminology at the time. The goal of Petit Mal was to produce a robotic artwork which is truly autonomous; which was nimble and had 'charm'; that sensed and explored architectural space and that pursued and reacted to people; that gave the impression of intelligence and had behavior which was neither anthropomorphic nor zoomorphic, but which was unique to its physical and electronic nature. It was not my intention to build an artificially intelligent device, but to build a device which gave the impression of being sentient, while employing the absolute minimum of mechanical hardware, sensors, code and computational power. My focus was on the robot as an actor in social space.

Although much work has been done in the field of screen-based interactive art, the 'bandwidth' of interaction in these works is confined by the limitations of the desktop computer. I am particularly interested in interaction which takes place in the space of the body, in which kinesthetic intelligences, rather than 'literary-imagistic' intelligences play a major part.

The formulation "autonomous robotic artwork" marks out a territory quite novel with respect to traditional artistic endevours as there is no canon of autonomous interactive esthetics. Petit Mal is an attempt to explore the aesthetics of machine behavior and interactive behavior in a real world setting. Petit Mal seeks to raise as issues the social and cultural implications of 'Artificial Life'. I wanted to avoid athropomorphism, zoomorphism or biomorphism. It seemed all too easy to imply sentience by capitalising on the suggestive potential of biomorphic elements such as eyes, ears, legs, arms etc. I did not want this 'free ride' on the experience of the viewer. I wanted to present the viewer with a phenomenon which was clearly sentient, while also being itself, a machine, not masqerading as a dog or a president.

I must emphasise that as an artist, I am an amateur and an impostor in fields of robotic engineering, artificial intelligence and cognitive science. My knowledge is unsystematic, it has been acquired on the basis of need and interest. However my outsider status has allowed me an an external and interdisciplinary perspective on research in these fields. It became clear during the project that the project had much in common with research into artificial life and autonomous agents. I have discovered that some of my basic assumptions align me with the more progressive research in those fields. Coming from an artistic

background, I have a strong sympathy for an 'embodied' approach. I also have an holistic approach to design and construction. I have designed and built most of the robot myself from basic components, often customising industrial surplus hardware.

Physiognomy and design theory

I wanted to build a device in which its physiognomy was determined by brutally expedient exploitation of minimal hardware. The two wheeled design offered the most expedient motor realisation for drive and steering: two pulse width modulated DC gearheard motors. This two wheeled design then demanded a low center of gravity to ensure stability. This swinging counterweight then demanded a solution to the problem of the stabilisation of the sensors so that they wouldn't swing radically, looking first at the ceiling then at the floor.The second internal pendulum gave this stability. In this way the structure specified the necessary extrapolations to itself, the development of the mechanical structure was not a gratuitous design but a highly constrained and rigorous engineering elaboration based on the first premise of two wheeled locomotion.

The double pendulum structure then implied a separation between logic and motor parts, the lower or outer pendulum carries motors, motor batttery and motor drive electronics, the inner pendulum carries the sensors at the top, the accelerometer in the middle and processor and power supplies as counterweight in the lower part. The batteries are not dead weight but in both cases also function as the major counterweights. The inner counterweight provides passive self stabilisation for the sensor head. It then became clear that the angle between the two pendulums could be measured and this angle could be used. The analogy to the semi-circular canals of the inner ear as the primary sensor of balance in humans is clear: the accelerometer is a rudimentary proprioceptive sensor, it measures relationships between parts of the robot's 'body'. It was important to me that this robot was 'aware' of its body.

From the outset I wanted to approach hardware and software, not as separate entities but as a whole. Data collection requirements necessitated the development of the stable inner pendulum, likewise the physical structure, together with the basic requirements of navigation and interaction with humans, determined the choice of sensors and the code. The suite of sensors is abolutetly minimal: three untrasonics, three pryo-electrics, two very low resolution encoders and an analog potentiometer on the 'accelerometer'. I wanted the software to 'emerge' from the hardware, from the bottom up, so to speak, The code would make maximal utilisation of minimal sensor data input. Petit Mal has had four sucessive sets of code, each increasingly more subtle in its adaptation to the dynamics of the device and more effectively exploiting the minimal processor power (one 68hc11).

The heart of the mechanical structure of the robot is a double pendulum, an inherently unpredictable mechanism. Emblematically, this mechanism stands for the generative principal that the machine, as a whole, is unpredictable, and a little 'out of control'. This is the logic behind the choice of name for the robot, in neurologial terminology, a Petit Mal is an epileptic condition, a short lapse of consciousness. The humour of this notion originates in the way in which it is contrary to the conventional idea of 'control' in robotics. Petit Mal has essentially no memory and lives 'in the moment'. The code, while not adhering to a subsumption model, has strong sympathies with bottom up approaches.

My approach has been that a cheap solution (in labor, money or time) to a particular problem which was 70% reliable was preferable to a solution which was 90% reliable but cost several times as much. Part of the rationalisation for this was that the very fallibility of the system would generate unpredictability, behavior, personality. It was pointed out to me by an engineer that my 'under-engineering' approach could lead to a much wider range of possible (though unreliable) solutions. The field of possibility is thereby expanded. Eventually such solutions could be refined. He was of the opinion that this approach could lead to better engineering solutions than an approach which was hindered by reliability in the research phase.

In robotics circles one hears the expression 'fix it in software'. This is applied to situations when the hardware is malfunctioning or limited. This expression is emblematic of a basic precept of computer science and robotics, the separation of hardware an software and the privileging of abstract over concrete. I reject this position. I have attempted, in Petit Mal, to forge an alternative to this dualistic structure. My approach has been that the limitations and quirks of the mechanical structure and the sensors are not problems to be overcome, but generators of variety, possibly even of 'personality'. I believe that a significant amount of the 'information' of which the behavior of the robot is constructed, is inherent in the hardware, not in the code.

My experience has shown that 'optimization' of the robots behavior results in a decrease in the behaviors which to an audience confer upon the device 'personality'. In sense then, my device is 'anti-optimised' in order to induce the maximum of personality. Nor is it a simple task to build a machine which malfunctions reliably, which teeters on the threshold between functioning and non-functioning. This is as exacting an engineering task as building a machine whose efficiency is maximised.

Many scientific ideas become dangerous when extrapolated into the social realm. One such is the idea of the univeral machine. Although in the abstract, one may concieve of a generic serial processing computer as a formless hardware vehicle which takes on a 'shape' determined by the currently running code, any real world machine, such as a robot, has unalterable physical characteristics which constrain and define its nature. My machine is a particularly eccentric mechanism. Many researchers in both the sciences and the arts apply this notion of 'universality' culturally and socially. They explicitly or implicitly treat the computer as a culturally 'neutral' object. But we all know that a computer or a robot are full of cultural meaning, they are emblems of the post-industrial world in both its utopian and distopian aspects.

Behavior, interaction, agency

People immediately ascribe vastly complex motivations and understandings to the Petit Mal. The robot does not possess these characteristics or capabilities, they are projected upon it by viewers. This is because viewers (necessarily) interpret the behavior of the robot in terms of their own life experience. In order to understand it, they bring to it their experience of dogs, cats, babies and other mobile interacting entities. The machine is ascribed complexities which it does not possess. This observation emphasises the culturally situated nature of the interaction. The vast amount of what is construed to be the 'knowledge of the robot' is in fact located in the cultural enviroment, is projected upon the robot by the viewer and is in no way contained in the robot.

Such observations, I believe, have deep ramifications for the building of agents. Firstly, any effective agent interface design project must be concerned with capitalising on the users' store of metaphors and associations. Agents only work only because they trigger associations in the user. So agent design must include the development of highly efficient triggers for certain desired human responses. An application of semiotics is required. In his painting 'Ceci n'est pas un pipe', René Magritte encapsulated the doubleness of symbols and the complexity of representation. This signification can be used to good effect in agent design: a very simple line drawing of a pipe, for instance, triggers a rich set of associations in the user. However, for the same reasons, these associations, like any interface, are neither universal nor intuitive, they are culturally and contextually specific.

Another curious quality of Petit Mal is that it trains the user. Due to the desire of the user to interact, to play; no tutorial, no user manual is necessary. People readily adopt a certain gait, a certain pace, in order to ellicit responses from the robot. Also unlike most computer-based machines, Petit Mal induces sociality amongst people.

Caucus

While developing Petit Mal I rejected the idea of a group of such robots: flocking was a solved problem. More recently it became possible to conceive of a technically feasible experiment in synthetic sociality. Hence, since 1995, I have been developing a project for a group of roughly human scale robots, based on Petit Mal, which interact with each other via the exchange of language tokens and individually generate sociality on-the-fly. Raw sensor data is not transmitted, nor is a paradigm of distributed processing embraced. Each robot maintains its individuality: in the process of reacting to its environment, broadcasts its 'opinion' of its situation via linguistic tokens. Each robot enacts behavior based not only on the data recieved from its sensors, but on its 'understanding' of the utterences of the other robots, and their relative locations.This project has become a focus of the Social Robotics Syndicate, formed in mid-1997, a professional colaboration between Kerstin Dautenhahn and Simon Penny.