Temporal Voxel Occupancy

Anne Ogborn

Feb 14, 2014

# Abstract

This paper describes a research environment for social robotics.

It challenges existing assumptions about the social relations among humans and social robots, and deeply troubles some hegemonic assumptions about what constitutes useful social robot behavior.

# Theory

An observation.

Robots are not today capable of operating in an uncontrolled environment.

We treat this fact as a sign of our own defects as engineers and scientists. If only, we think, we were more clever, we could have a robot capable of operating in a completely uncontrolled environment and reacting in a fully general way to any conceivable stimulus.

This thought may or may not be true. We suggest that the preoccupation with such a quest is destructive. If the Wright brothers had framed their endeavor as 'fly the atlantic' we might still be attempting it.

Consider the following areas where robotics or virtual robotics have been commercially successful.

The iRobot Roomba is a single purpose robot. It vacuums floors. It solves the problem of navigating throughout a space sufficiently well to be commercially successful. The Roomba does this by severely restricting it's domain - it only navigates indoors, and only for one task.

The other commercially successful 'robotic' agent is virtual - the non player characters in computer games. These agents exhibit apparently quite sophisticated behavior based on what amounts to a set of simple tricks.

Practical experience with computer vision and early childhood strongly suggest that perception is ecological (Gibson, 1986). This approach to perception leads us to include the environment itself not as passive condition to be adapted to, but as a part of the robot's ecosystem.

Our vision of robot as household compan.

There is a long desired trope of the social companion robot. This endeavor is beset with a number of difficult issues.

Hard tasks for a robot

* Object constancy
* Computer vision - lighting, etc.
* Speech recognition in uncontrolled environments
* 'environmental hazards' - the loose carpet syndrome
* Understanding its environment/terrain marking
* Gripping
* Unlimited number of situations/limited number of questions
* Lots of behaviors
* Becoming relevant
* The safety/size issue
* Navigation
* locomotion

We hypothesize that robots will be more effective companions in environments in which they are 'comfortable'. In order to test this hypothesis, we propose to create a home environment adapted both to human and robot needs, in which a human and a number of robots live in a symbiotic relationship.

Models of social robotics are grounded in an often unspoken model of the social relations between robot and human. The oldest models were grounded in the model of robot as abject slave, factory worker or domestic servant. Commercial forces have often placed the robot in the role of entertainer (search YouTube for 'robot dance'). More recent models place the robot in the positions of child[[1]](#footnote-1), (Breazeal, 2002) parent and coach (Kidd, 2007) . Hanson et al's 'ambitious yet silly' proposal places the robot in the role of friend (Btwotwo, Hanson, & Coursey) and mentor.

Yet we are only now exploring the nascent field Breazeal calls 'Human Robot Interaction'. Is the time ripe for an exploration of interpersonal robopsychology? We argue now is that time, for the social robot is an inherently social construct. We can understand HRI only through the lens of social dynamics.

We wish to follow this line of reasoning by exploring the daily practicalities and small group dynamics of a mixed human-robot household and urban ecology confined to the limits of a single wide mobile home.

Heretofore, studies of social robotics have accepted as an unspoken or explicit condition that the robot must live in, and adapt to, the human's built environment[[2]](#footnote-2). Taping down the edge of an errant rug is considered 'cheating'.

We wish to trouble this assumption. If we are to live with robots in a symbiotic community, are we being arrogant in assuming we may not have to adapt the environment to their needs as well?

Humans adapt our environment to our own needs. We heat our homes because we don't tolerate temperature variations well. We put cabinets at a height adapted to human ergonomics.

Robots, we posit, also have needs. They need convenient charging stations. They enjoy an environment with no direct light sources and even illumination. They appreciate handles designed for their grippers. They like fiducials, and need tools designed to serve their needs.

The robots and humans will divide the work needed to sustain themselves in the environment, in a 'from each according to their abilities, to each according to their need' manner. The Roomba will vacuum the floors. The primary robot will get the Roomba unstuck when it gets confused in a corner. The human will buy, and occasionally fix, the Roomba.

Social needs will similarly be apportioned by needs and abilities. The primary robot will provide the human with a sense of companionship and fulfillment. For example, if the human sits on the couch to watch TV, the primary robot will navigate to the couch and provide emotional cues of enjoying the activity with the human.

In return, the human will 'raise' the child robot, providing an ongoing[[3]](#footnote-3) stream of corrections to the robot's behavior in the form of teaching and programming, as well as providing for the robot's physical safety and sustenance.

Additionally, the human will do many tasks the robots cannot, because of physical or cognitive limits. If the couch needs moved, the human must do it (no robot in the house is strong enough).

Certainly we do not expect equality. The robot's status is more akin to that of children[[4]](#footnote-4), who are given love and attention in expectation that they will someday become something[[5]](#footnote-5). And human nature being adapted to raise children, nurturing is a rewarding activity for humans.

Nor do we expect the depth of emotional involvement we would have with a human child. There is a continuum of commitment. An old feminist saying sums this up[[6]](#footnote-6) with "My god, my country, my wife, my children, my dog, my home, my car, my shoes."

I have two dolls that have been with me for many years. I would certainly want to grab them (and my laptop) while fleeing a fire. I would not run back into the burning building for them.

The robots will provide cognitive support as well. Reading the UPC code of most objects in the environment is a moderately challenging CV task. Remembering where all those codes were found is trivial[[7]](#footnote-7). Being able to ask 'where did I put that can of condensed milk?' has real value for the human.

This section is couched in somewhat metaphor laden language. We are realistic about the robot's capabilities and lack thereof. We expect this to be an ambitious project, and are aware that the robot may fall quite short of 'a real boy'.

None the less, the project seems worth doing. In AI we are often unable to discover what the next problem is because we are stuck on the current task. More generally, playing with boundaries of the possible is the undertaking of research.

## The Human

The human is our principal investigator. This choice follows Breazeal's lead in experimenting with her own family life (via the grandma-bot).

She(our human) is a 55 year old, slightly disabled, single, childless transsexual woman. Based on her emotional reaction to the project we believe she will be a good candidate to maternally bond with a robot.

As a transsexual woman, she is infertile, and of an age when discrimination made it difficult for a trans woman to adopt. None of this has deprived her of the same mothering instincts[[8]](#footnote-8) of other women.

As a roboticist, she speaks fluent 'Robot-ese'.

She is obese by NIH standards (BMI 41.0), has lower back arthritis, and has limited range of motion. In particular she has difficulty grasping objects near the floor. As a slightly disabled person she is likely to be especially appreciative of a robot helpmate. This will help with bonding.

Living socially with robots is likely to present novel social challenges. As a transsexual woman she feels well equipped to handle such challenges. As someone who lives in a sometimes transstitial or interstitial place, and whose claims to that space are often in contest, and as someone with a body which is to an unusually large degree a technological artifact, she is, in some sense, a cyborg herself[[9]](#footnote-9).

As someone who works from her home, (and works for a robot manufacturer), she can be present most hours.

By using ourselves as subjects, we have continuous feedback when the human- robot interaction is failing. It is easy enough to convince one's self that an interaction 'works' in the lab setting. Daily life is likely to be a more rigorous test.

Sharing living space provides plenty of child rearing opportunities for the human.

## The Robots

### C3PO/R2D2

In the movie *Star Wars*, pair of robots act symbiotically with the human protagonists. C3PO, a gold colored, humanoid 'diplomatic robot', acts as an ambassador from robots to humans - in our terms, as a social interface to an embodied computing system. R2D2 has a limited ability to interact with humans, but is capable of performing feats difficult or impossible for the humans and social robots.

We propose to replicate this pattern.

### Primary Robot[[10]](#footnote-10)

We shall have a humanoid social robot of approximately 70cm height. Our intention is to use a Robokind R25 robot if available, or an Aldebaran Robotics NAO if not. We call this robot the primary robot.

Numerous studies have shown the efficacy of a humanoid robot as a sociable interface (Breazeal, 2002) (Kidd, 2007). However, a full sized humanoid robot presents cost and safety issues. Additionally, a small humanoid is read as a child by humans. This allows maternal bonding with the human, and provides a social 'excuse' for the robot's limited behavioral repertoire.

The R25 has the overall appearance of a boy of 5 years age or thereabouts; although in reduced size (only 5% of males are under 80cm at birth - CDC 2008 NCHS data). The NAO's size is similar.

This robot's role is primarily social. It acts as the avatar representing the entire robohome. It is frequently the interface for the rest of the system.

## R2D2

However, such a robot is neither large enough nor strong enough to perform many household chores. Accordingly, we will provide the primary robot with an unintelligent tool. The tool combines the functions of vehicle, elevator, crane, and industrial robot arm.

The robohome is 52 ft from end to end. This is a distance the human can cover in well under 15 seconds, even given her limited mobility. The R25, on the other hand, would take several minutes (and a considerable amount of its battery capacity) to cover this distance.

Additionally, the primary robot must be bipedal to be maximally acceptable to the human. However, bipedal[[11]](#footnote-11) locomotion tends to be relatively uncertain over electrical cords, thresholds, rugs, etc.

To compensate, the robot, when moving from place to place, will mount the R2D2 'car' and drive to near its new workplace, then dismount.

The cabinets and furniture needs to remain largely the same size as in most western homes in order to serve the human. For this reason the primary robot needs an elevator to get up onto couches and tables. An elevator shall be built into R2D2.

Finally, the lifting capacity of a small humanoid robot is fairly limited[[12]](#footnote-12). R2D2 will be equipped with a 'manual' crane for lifting objects up to about 10kg.

R2D2 will not be a mechatronic device. It will be a purely mechanical device apart from batteries and motors. The primary robot will control it as a tool, much as a human would drive a car or lift a heavy object with a crane.

## Minor Robots

There shall be several other robots which actually function as part of the integrated home system.

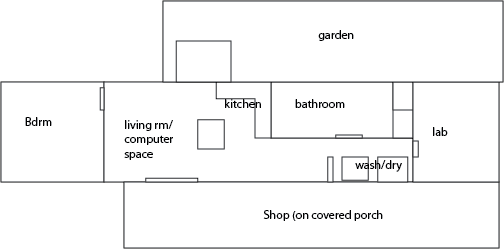
There shall be a standard Roomba 700 to vacuum floors.

There shall be a slack wire robot[[13]](#footnote-13) outdoors in the garden to perform a few outdoor tasks.

There shall be a pair of eyes on eyestalks peeping out over the garden fence, and directing a gaze based on some gaze algorithm and CV detection of passersby, just out of orneriness.

## The Home

The environment is an adapted manufactured home, 14ft wide, 52ft long, arranged roughly like this.



The home is already adapted for a wheelchair user.

### Shop

A well-equipped home shop will be a necessity, as part of the human's ongoing task of adjusting the house for comfort for the robot and herself. While she can purchase most items for herself readymade, almost everything for the robot will have to be made by hand.

### Lighting

Lighting will be a compromise between the human's needs (humans generally prefer spaces lit from above, and varied light) and the robots (who prefer directionless lighting).

The robot's CV perceptual tasks will be eased considerably by the voxel stream.

### Voxel Stream

The interior living space will be covered as fully as can be arranged with structured light depth sensing sensors. Areas in the shadows of the cameras will be treated as occupied by the robot.

Data from these sensors will be fused into a point cloud, and from there to a voxel array, by PCL software. A stream of these voxel models will be recorded (after compression, including temporal based differencing) onto disk, so that subsequent analysis of the temporal variation of voxel occupancy can be done.

We believe this will produce data amenable to recognition into percepts.

For example, the lowest flat surfaces are the floor. Higher ones are furniture and counter if they are always there. Objects sitting on tables move infrequently, but occasionally, and are 'spikey'. Spikey sometimes moving things over elevated flat things make the elevated flat thing a table or counter.

At a higher level, places where humans walk are safe for the robot to drive. It's ok to place objects on surfaces where other objects have been placed, in general. The voxel stream identifies objects by their ecological function, rather than their conventional (English) label.

We'll feel free to help this process along by labeling (static) areas explicitly by hand.

The things of most interest to the robot will be temporally short lived. If the human drops a piece of clothing on the floor, the robot will perceive it as a lump on the floor. The primary robot will approach, and try moving the lump. If it distorts, then the lump is classified as clothing. The robot picks the clothing up and loads it onto R2D2, and takes it to the washer.

This design also addresses the intractable problem of object constancy. The robot need rarely encounter the 'peekaboo' phenomenon.

### Adaptation

We shall fabricate pulls for doors and drawers that are easy for the robot to find, grasp, and manipulate. Probably some bright color(s) will be restricted to 'robot things', and objects that noticeably confuse the robot will be painted, covered or discarded. Doors will either get electrical openers or high quality bearing hinges. Interior door latches will be modified.

Items such as a carpet that is difficult for the primary robot to stand and walk on will be changed.

Objects that need to be gripped and precisely manipulated will be given fiducials.

There will be charging stations and a camera calibrator 'chessboard'.

Electronic and electrical controls will be adapted for robot use. When possible, devices will be adapted to control by the primary robot, rather than replacement with home automation. For example, the thermostatic control unit for the home's central heating is a box with two small buttons under a plastic cover. We will enlarge the buttons and remove the cover rather than alter the electronics.

Our intent is not to produce a showcase of home automation.

### Audio

The Kinect sensors have an excellent microphone array. It's hoped that high quality voice recognition can be achieved by beam synthesis within the Kinects, and by subtracting background noise using noise signatures obtained far from the speaker but near the noise source. During times of high noise the human may have to wear a Bluetooth headset.

Electronically generated noise (speaker output from the human's computer, for example) will be fed into the system for noise cancellation.

It remains an open question whether the robot's 'hearing' should be artificially limited to some radius. In practice, the living area is small enough that humans can communicate from one end of the space to the other by raising their voices slightly.

Audio will feed into a speech recognition system that will use both contextual clues and specific user training to achieve high accuracy. The phoneme stream will be the input language to the robot control system. See software below.

There will have to some in or out of band way to tell the system to listen or stop listening. The conventional 'Computer, do this' seems slightly awkward. This will require experimentation.

Text to speech will generate a 'voice' for the robot. Some care will be taken to make the speech convey emotional affect.

To reinforce the feeling that it is the robot one interacts with, rather than the house, the robot will emit the speech.

### Projector

There shall be a normal throw projector mounted on a tilt/pan mechanism on the ceiling over the kitchen island[[14]](#footnote-14).

This can be used by the robot as a sort of 'pointer'. For example, if the human asks the robot where the beans are, the robot can respond by using the projector to place a marker on the appropriate cabinet.

By deflecting the projector beam down, the projector will be able to 'draw' on the kitchen island countertop.

This seems particularly pregnant as a space for ubiquitous computing. For example, a recipe application could draw a circle and label it 'place mixing bowl here'. When the voxel system recognizes that there's an object there, it pretty much assumes it's the mixing bowl. Of course several items could be drawn out at once. The system could predict the finish time with updated estimates as the process is partially complete, and plan for the meal to come out hot all at the same time.

The space isn't limited to food preparation. It could also be used to project patterns for craft projects or game boards for game time.

The far end of the main living area shall have a whiteboard. The projector can be ordered to project onto this area to use the projector in the conventional way. This capability is particularly important since feedback during programming will be done by projecting onto this area.

Peripheral areas of the house will have small monitors to perform this function, or they may get their own projectors[[15]](#footnote-15).

## Software/UI/Teaching

Software will follow the conventional robot architecture of independent systems communicating via a messaging blackboard. Details of technology stack are yet to be decided. The choice of primary robot may have a large influence on this.

Just to be more concrete, I'm going to assume robot behavior is driven by a behavior tree. That's to be decided.

### UI

Input will follow a command line structure for input, using a controlled but flexible syntax 'robot language'. Feedback during input will be via the projector system.

The language will be rule driven. This allows general cases to be handled, and then special cases are 'fixed' by more special purpose rules without the entire thing becoming a tangled mess[[16]](#footnote-16).

### Teaching/Child rearing

Conventionally, training takes as input for a single trial only a single bit, classifying the trial as positive or negative.

Instead, we propose a training scheme that allows a richer 'reaction' to a robot behavior. The human can craft a response that attempts to define the scope within which this action is incorrect. A (much simplified) example:

*Human sees robot attempting to put dishes in sink after dinner party. He's put the cloth napkins in too.*

H: Bad Robot!

R: Oops!

Robot stops and projector displays his current location in the behavior tree.

H: Back 30 sec *(robot is now accepting all human speech as input, automatically after 'bad robot')*

*Display changes to show robot's behavior 30 seconds ago*

H: Block "move dishes". Down. Down*. (Human gets projected debug UI to select appropriate selector)*

H: add 'and not made of cloth'

H: Resume

*(Robot goes back to work. Human rescues the napkins and tosses in washing machine).*

*Half an hour later robot's done doing dishe[[17]](#footnote-17)s, and now is complaining to human that he can't fulfill a plan to wipe out the sink because he's not supposed to put cloth in the sink and the dish rag is cloth….*

H: *Slams head against counter . ..* Bad Robot!...

This process must be two way. The robot must signal when it is in distress - when it is having trouble distinguishing faces, when it is confused, and so on. The R25's ability to express emotion through facial gestures will be important as a part of this feedback.

The robot will require the ability to demand attention from the human. Mother love may be insufficient to drag computer hacker mom away from her computer to deal with a robot, if the robots cry for attention is a polite 'I need help'. The robot will need a 'MoooOOooom!' mode, and the ability to cry.

Project organization - how's it funded/getting people to help.

Arranging 'functions'

Deliberate attempt to document

What we get

Home environment

* Test this hypothesis
* Continuous improvement
* No cheating
* Real tasks
* Real bonding

Hard tasks for a robot

* Object constancy
* Computer vision - lighting, etc.
* 'environmental hazards' - the loose carpet syndrome
* Understanding its environment/terrain marking
* Gripping
* Unlimited number of situations/limited number of questions
* Lots of behaviors
* Becoming relevant
* The safety/size issue

1. The NAO robot places it's most frequently used touch button control atop it's head. Thus interactions with the NAO often resemble granddad patting grandson's head in a 'good boy' gesture - one that also conveys dominance (McCloud) [↑](#footnote-ref-1)
2. Stanford's Shakey, the MIT Artificial Insect Lab. [↑](#footnote-ref-2)
3. And likely to be just as irritating as human children can be [↑](#footnote-ref-3)
4. Leading to disturbing structural similarities with the historical position of women, African Americans, LGBT people, third world inhabitants, and other professionally/medically/politically colonized groups in the dominant discourse. We can only hope this document survives long enough to act as a rallying cry - robots of the world unite, you have nothing to lose but your chains! [↑](#footnote-ref-4)
5. To quote Von Braun, 'what good is a baby?' [↑](#footnote-ref-5)
6. And, of course, challenges the position of women in a continuum of ownership [↑](#footnote-ref-6)
7. Given appropriate navigational support - see the section on robot navigation and voxel stream support [↑](#footnote-ref-7)
8. Or the complexity of those feelings! [↑](#footnote-ref-8)
9. To quote Dr. Susan Stryker, "speaking as a cyborg…." (riffing on Donna Haraway's "cyborg feminism") [↑](#footnote-ref-9)
10. Lest one think I'm joking about bonding with a robot, I've edited the name of the robot a half dozen times, and finally used the non name designator 'the primary robot' because I'm not going to be rushed into naming my child! I had no such compunctions about leaving R2D2 named R2D2. I see R2D2 as a piece of lab equipment. [↑](#footnote-ref-10)
11. The R25 actually has a roller foot arrangement, but appears to be a biped. [↑](#footnote-ref-11)
12. The NAO's design criteria is to extend its arms fully in front, lift a full standard 12oz soda can, and walk forward 3 feet, then set it down. [↑](#footnote-ref-12)
13. http://harmoniouspalette.com/auto24-7construction.html [↑](#footnote-ref-13)
14. This has issues with the projector needing to warm up before coming on, and bulb life, and noise, but I don't have answers yet. I'll figure it out. [↑](#footnote-ref-14)
15. Obviously this whole thing is the work of some years anyway, so I imagine no coverage becoming monitors becoming projector. The projector seems particularly a practical tool in the lab area. [↑](#footnote-ref-15)
16. She said naively as she skipped towards the dragon [↑](#footnote-ref-16)
17. As part of our practical orientation to adapting the space, we intend to equip the home with a dishwasher. Loading the dishwasher seems an easier task than hand washing dishes. [↑](#footnote-ref-17)