DELIVERABLE 3.4
Final evaluation of ACCOMPANY computational memory architecture

Author(s): *Joe Saunders*

Joan Saez

Dag Syrdal

Kerstin Dautenhahn

Project no: *287624*

Project acronym: *ACCOMPANY*

Project title: *Acceptable roboticsCOMPanions for AgeiNg Years*

7th FRAMEWORK PROGRAMME

ICT Call 7 - Objective 5.4 for Ageing & Wellbeing

Doc. Status (draft, proposal, …)	final

Doc. Nature (report, …):	Report
Version	5.0
Actual date of delivery	October 1, 2014
Contractual date of delivery	M36
Project start date	October 1, 2011
Project duration	36 months
Approver	UVA

DOCUMENT HISTORY

Version	Date	Status	Changes	Author(s)
1.0	**2014-08-28**	**Draft**	**First version**	**Joe Saunders**
2.0	**2014-09-23**	**Draft**	**Edited version**	**Kerstin Dautenhahn**
3.0	**2014-09-25**	**Draft**	**Version 3**	**Joe Saunders**
4.0	**2014-09-30**	**Final**	**Version 4**	**Joe Saunders**
5.0	**2014-10-21**	**Final**	**Version 5**	**Joe Saunders**

AUTHORS & CONTRIBUTERS

Partner Acronym	Partner Full Name	Person
UH	**University of Hertfordshire**	**Joe Saunders, Joan Saez, Dag Syrdal, Kerstin Dautenhahn**
Reviewers:	**University of Amsterdam**	**Ninghan Hu**

Table of Contents

SHORT REPORT

This document describes the final evaluation of the Accompany computational memory architecture described in deliverables D3.2 and D3.3 and implemented on the Care-O-Bot3®, the companion robot in the ACCOMPANY project.

We report on two areas of *formative* evaluation of the architecture relevant to the aims of Accompany. The first is the evaluation of the robot teaching and learning components and the second is the evaluation of the visual episodic memory component. The design and implementation of the robot *teaching* components were described in detail in D3.3 and will briefly be summarised in this document, however the design and implementation of the *learning* components will be described in detail together with their formative evaluation. The visualisation component was also described in D3.3 and will briefly be recapped here together with the results of the formative study. Note that the complete memory architecture has also been formally evaluated as part of the long-term summative studies reported in D6.8.

One of the aims of the ACCOMPANY project is to provide aid and support for elderly people and allow them to reside for longer in their own homes by providing facilities not only for care, but also for companionship. It is envisaged that companionship would be supported by the dual functions of *co-learning* and *re-ablement*.

This document reports on 1) A recap of the background to this research and a definition *of re-ablement and co-learning*, 2) A brief reminder and explanation of the architecture itself, together with a general description of the environment, the robot and our ontological approach, 3) A brief summary of the teaching component and a description of the formative studies and results, 4) A brief summary of the memory visualisation component and a description and results of the formative studies carried out, 5) A detailed description of the new learning components and details of an initial formative evaluation with an informal carer group, 6) A final round-up and discussion of achievements and possible future directions for this research including some preliminary work on sequential action prediction (described in the appendix).

Contents

1 Introduction

The memory model used for both the ACCOMPANY summative and formative studies was described in detail in the previous deliverable (D3.3). This model supports robot self-learning and adaptation in order to cope with supporting user-robot interaction in a multitude of physical, cognitive, social and communicative tasks. Additionally, visualisation of the robot's memory contents allows users to understand the interaction histories with the robot – the aim of visualization being to provide transparency contributing to the user's trust in the system. By allowing the review of tasks carried out in the past, users have the opportunity to reflect on their personal daily routines. Past experiences stored in the memory can also be "re-lived" for social and narrative purpose when users want to share them with others (e.g. family members and carer) to create conversational topics.

In this deliverable we present results of a formative study carried out at the University of Hertfordshire Robot House that was used to evaluate the models above and gauge users' reactions regarding usage and usability. This formative study considers both the memory model and how users could individually personalize it using the robot teaching facility as well as the visualisation component.

These components we also evaluated in a long-term summative evaluation, and this is described more fully in deliverable D6.8.

One of the aims of the ACCOMPANY project was to provide aid and support for elderly people and allow them to reside for longer in their own homes by providing facilities not only for care, but also for companionship. It is envisaged that companionship would be supported by the dual functions of *co-learning* and *re-ablement*. In previous deliverables (D3.2/D3.3) we specified how our computational memory architecture was planned to support episodic, procedural and semantic aspects of memory and described their technical implementation. One important aspect of this work was that of robot teaching and robot learning. In this deliverable we also additionally report on work carried out to support robot learning and also on a small evaluation study of the robot learning system carried out with a subset of the user group (the informal carer group).

In order to provide details of these evaluations we divide the report into a number of sections describing: firstly, the background to this research and a brief recap of the technology, secondly, the formative evaluation procedure for both the robot teaching component and the memory visualisation, thirdly the results of the evaluation. A description of the learning component and its background research are then described followed by results from the initial evaluation study carried out to assess this aspect. Finally, some general conclusions and outlook for future studies is described as well as some preliminary work on sequential action prediction (described in the appendix) supporting these future objectives.

2 Background

It is predicted that many countries worldwide are facing a demographics problem over the following decades. This is due to increasing life expectancy, leading to more elderly persons, combined with a decrease in the proportion of younger people providing support to the elderly. Robotic companions have been suggested as an assistive technology to meet an ever ageing population. Our approach utilises a commercially available robot, the Care-O-bot3®, manufactured by Fraunhofer IPA (Reiser et al. 2009) sited in a fully sensorised house (a smart home which we call the *robot house*). The house itself is a typical British three bedroom house near the University of Hertfordshire and it was specifically chosen to be a realistic home environment rather than a modified scientific laboratory.

In order to achieve these aims the Computational Memory Architecture (CMA) was designed by ensuring the disciplined integration of the robot house sensor network, the sensory capabilities of the robot itself, and the social memory aspects from the user themselves, into a common framework. Additionally, a mechanism which allowed activities within the house, at both a sensory level and a more abstract contextual level to be joined as propositions (held as rules or preconditions), and optionally applied with temporal constraints, for resulting robot behaviours. Finally, the ability was added to be flexible in behaviour creation and scheduling. Given that our robot may be asked to carry out a large number of tasks, many of which may not be originally envisaged by the system designer, a flexible and ‘easy to use’ way of creating and personalising robot behaviours together with a mechanism for effectively scheduling such behaviours was required. Our goal was to make such facilities available to non-technical personnel such as the elderly persons themselves, carers or relatives.

3 Design and Implementation of the Computational Memory Architecture

In this work, behaviours are not pre-programmed, but instead *taught* to the system via one of two GUI interfaces (Saunders et al. 2013). The first GUI is a semi-technical interface allowing for direct access to all types of sensors and all types of robot actions, the second interface is more restricted and generates behaviours based on templates (see deliverable 3.3). This latter GUI trades behavioural generation complexity and expressiveness against ease of use and it is the interface which we would expect to be used by end users and the one which is formatively evaluated and the results discussed in this document. Note that the semi-technical interface was used for all other studies on the robot, including the summative study described in D6.8 and for scenarios 1, 2 and 3.

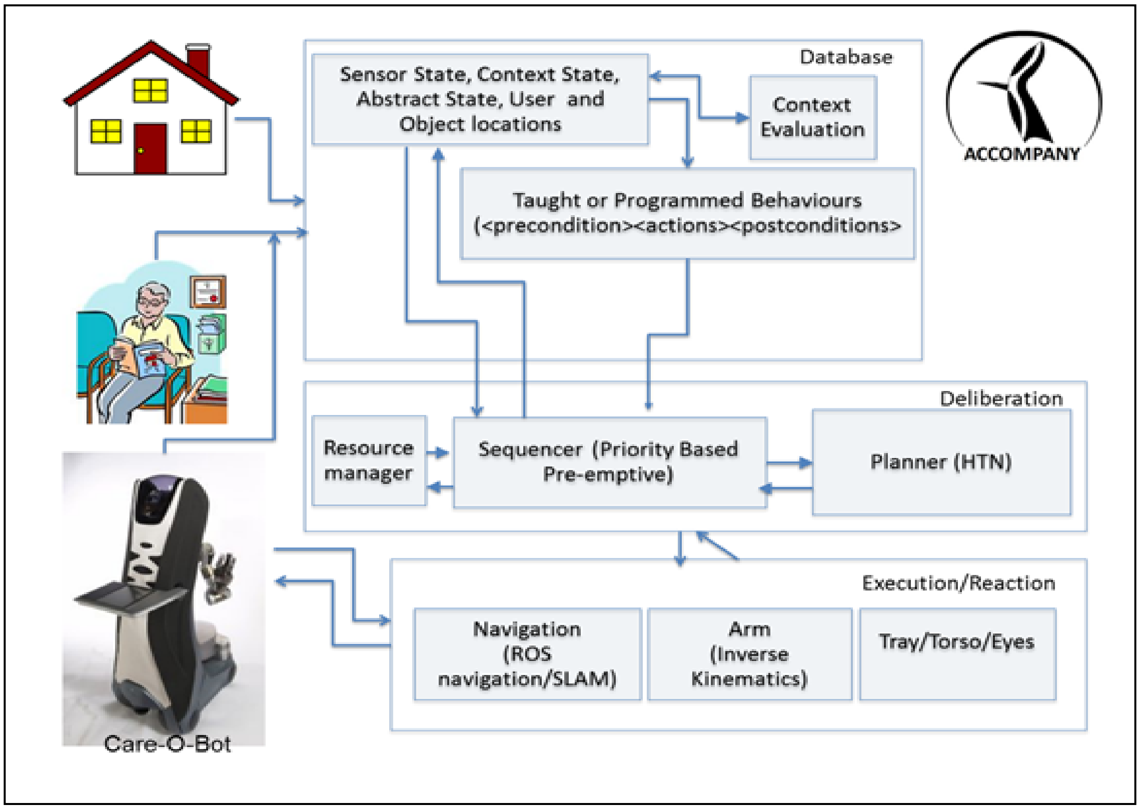


Figure 1: : A high level view of the control architecture for the Care-O-bot R within the robot house. The architecture is based on a typical three-layer approach comprising of deliberation (sense/plan), execution and reaction (act). The deliberative layer makes decisions based on real-time information from the central database. This information is composed of updates from the physical house sensors and updates made via the deliberation or context awareness rule-based programs to context or predicate sensors. The reactive layer operates primarily in a tight control loop in areas such as navigation and arm kinematics.

3.1 Brief Recap of Behaviour Creation and Execution

Behaviours that are generated follow the familiar pattern (similar to Nilsson's T-R formalism) of evaluating propositions (as *pre-conditions*), followed by execution of robot *actions* and updating of *post-conditions* (in a similar way to the add/delete lists of planning systems). Pre-conditions can evaluate any form of sensory information including both real sensors (set by environmental changes) and abstract sensors (context/predicate sensors). Post-condition updates are used to change predicate sensors; however physical sensors cannot be directly updated. An example of behaviour would be as follows:

BEHAVIOUR: doorbell

IF the doorbell has rung in the last 20 seconds (sensor pre-condition)

AND the user has not yet been informed (predicate sensor pre-condition)

THEN

send the robot to the user location in the house (action)

make the robot say 'Someone at the door' (action)

update the database to signal that

the user has been informed (update predicate sensor post-condition)

Example 1. A typical behaviour

If the set of preconditions are true, then the actions are executed, including the post-condition update. Actions, as well as being used to control the robot, can also be used to execute other behaviours, which again can have pre-conditions and actions. Sequential sets of actions can also be called, and these would be set-up as behaviours with no pre-conditions. Careful arrangement of behaviours allows the action of post-condition setting to fire other behaviours and thus 'fill in' details of the overall plan.

3.2 Additional Modules for Activity Recognition

Part of the work presented in this document (see section 6) is concerned with learning the activities of the house resident. The part of the overall architecture responsible for this is the "activity awareness" module (part of the "context awareness module" in Figure 1). This relies on a vector of sensory information (effectively the concatenated set of house/abstract/location and contextual sensors) polling a set of rules previously derived via the C4.5 rule induction system (Quinlan (1993)]. This module then updates abstract sensors associated with this activity. For example, an abstract sensor called 'watching TV' (created via the learning process described in section 6) might be set 'active' when the house resident was sitting at the sofa and the TV was on.

3.3 Robot Teaching

One of our goals in this work was to allow co-learning, whereby robot and user work together to achieve a goal. To achieve this we approach the problem in two ways. Firstly, by directly teaching the robot, via a GUI, what has to be achieved and when it should happen and secondly, via the robot (or the house) learning what is happening and learning to adapt to those conditions. The *learning* mechanisms are described in section 6.1. In the following sections, we describe the *teaching* interface used in the formative study. In order to create behaviours the user as a minimum would need to specify what needs to happen (the actions of the robot) and when those actions should take place (setting pre-conditions).

Take, for example, a user who wanted to be reminded to take their medicine at 5pm. In this example, the user need only specify what is important (that the reminder be given at a particular time) and the background system automatically generates the appropriate predicate pre-conditions to ensure the behaviour is generated accordingly.

This is possible as the majority of behaviours envisaged tend to follow common templates, and we exploit these templates to generate the appropriate conditional logic.

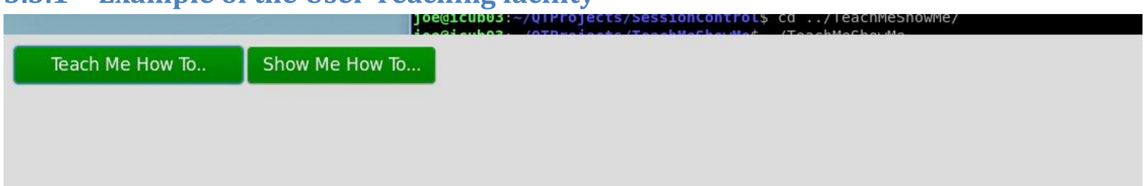
In this manner much of the cognitive load is removed from the user and left to the behaviour generation system. Co-learning is operationalized by allowing the robot to provide details of its existing sets of skills that can then be exploited by the user.

Thus the requirements of the user would be straightforward – WHAT to do – remind me to take medicine, WHEN to do it – at 5pm.

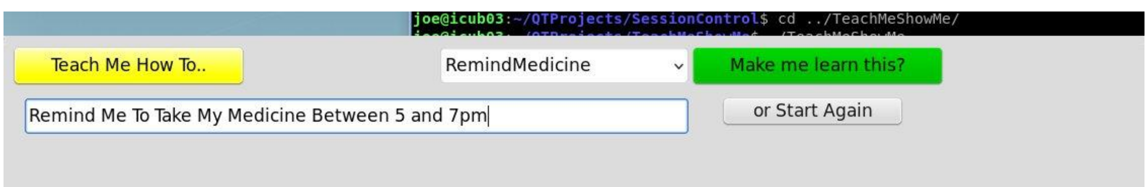
The underlying conditional logic can all be automatically generated (see deliverable D3.3 for a detailed explanation of this process).

The following section shows how the user enters information into the userGUI to create behaviours.

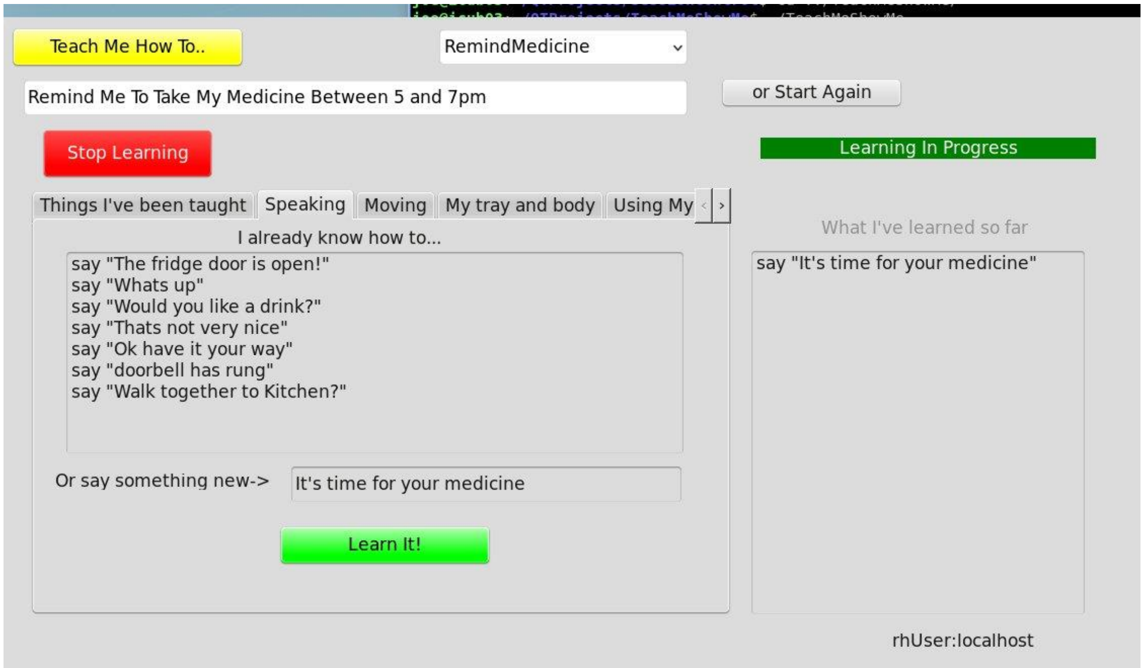
3.3.1 Example of the User Teaching facility



User is presented which the “TeachMe-Show Me” screen and presses “Teach me”.

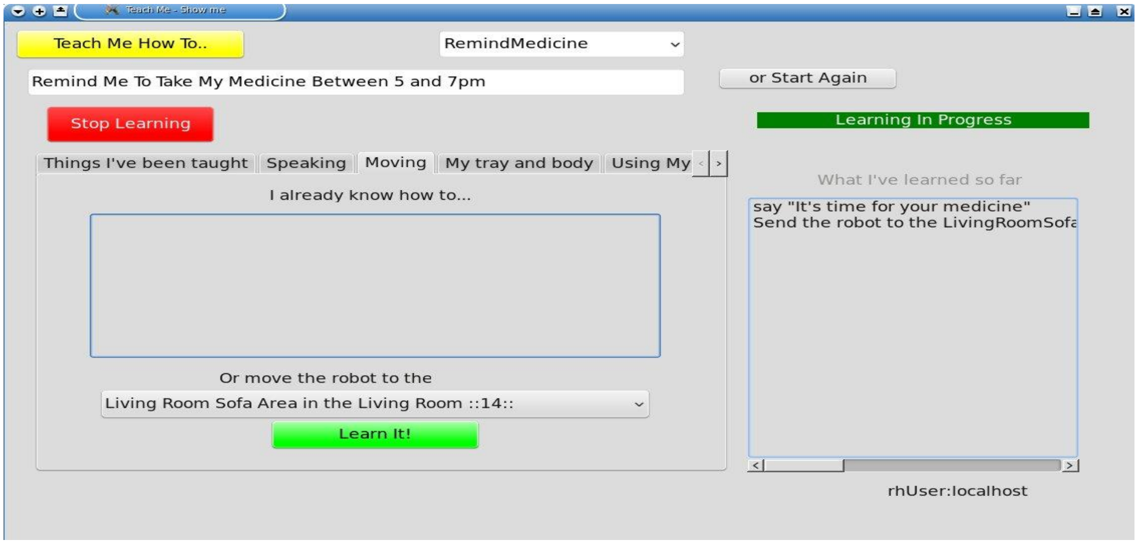


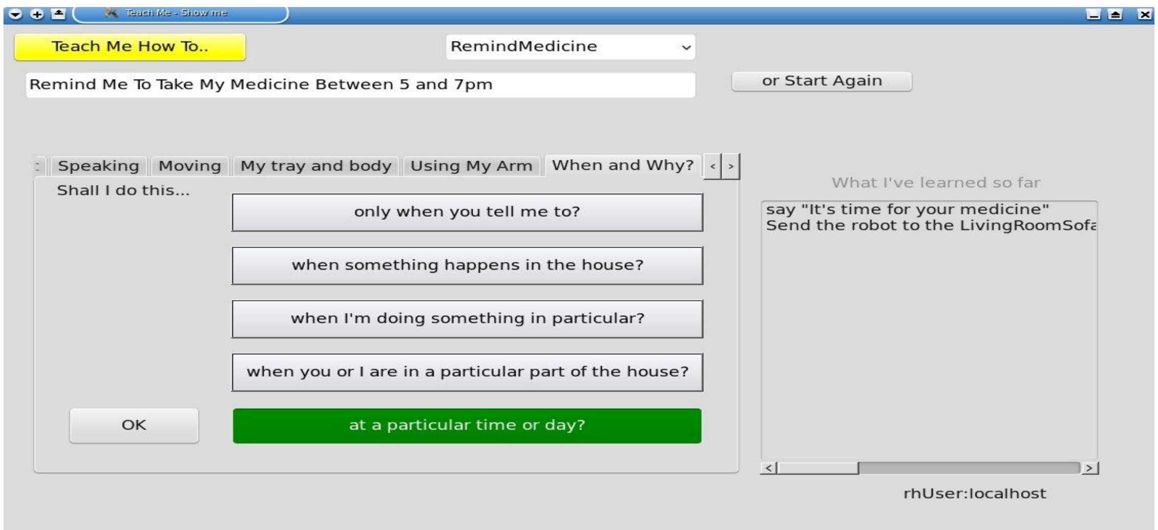
The user enters a text description of their requirement and clicks the “make me learn this” button.



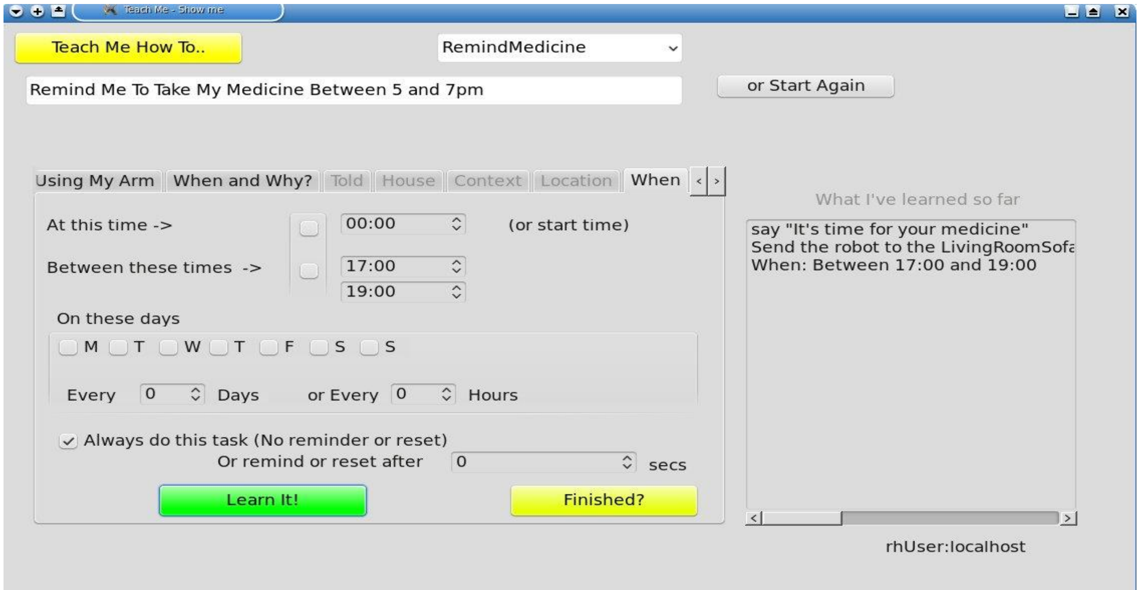
The “WHAT to do” screen appears. The tabs show the set of actions the user can teach the robot, and also shows what the *robot already knows* (the “I already know how to...” box). Clicking the “learn it” button causes the action to be saved into the “behaviour box” on the right.

The user clicks the “Moving” tab. The robot has not been previously taught this and so the “I already know how to do this...” box is empty. The user chooses to select a known position from the drop down box at the bottom. This gets added to the behaviour screen on the right.

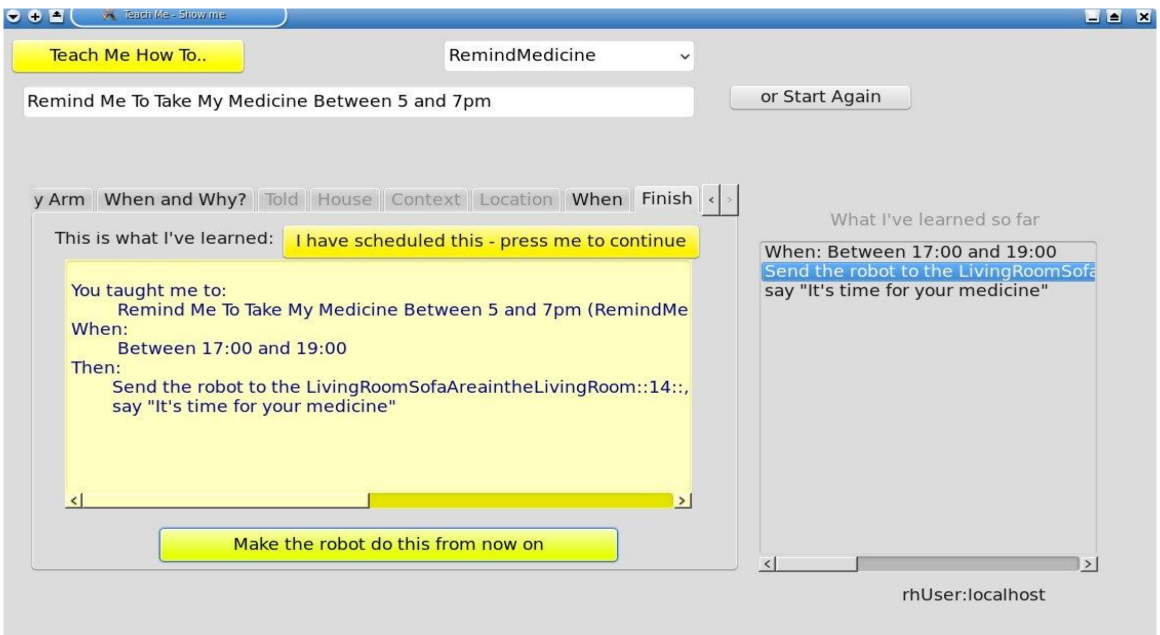




The “WHEN To Do IT” screen appears, and the “particular time of day” option is chosen. Note that multiple options can be chosen at this point.



Because the user chose the WHEN option as a particular time of day, a diary function appears. The appropriate times are entered. The user then clicks the "finished" button.



The complete behaviour is displayed to the user. They have the opportunity to go back and change it. Once finished the behaviour is sent to the scheduler.

4 Formative Evaluation of the Robot Teaching System

Evaluation of the template based teaching system was carried out in late February 2014 in the robot house. This involved 20 participants recruited from the general population. The summary procedure involved the following steps: 1) Participant completes consent form, personality and demographics questionnaires, 2) Participant is introduced to experimenter and robot, 3) Robot teaching evaluation – including questionnaires, 4) Visualisation Interface evaluation - including questionnaires, 5) Final questionnaires.

Our research questions for the formative study included not only usability of the teaching system but also how this varied with factors such as age, personality traits and attitudes to robots in the home.

4.1 Procedure

Each participant was introduced to the experimenter, a technician and the experiment psychologist. The technician was present only to ensure the safety of the participant (this is a safety requirement of the ethics approval for this study) and played no other part in the experiment. The technician was stationed in a part of the room outside the main interaction area.

The psychologist asked the participant to fill in a consent form, demographic, computer, robot experience and the Ten Item Personality Inventory (TIPI) (Gosling et al. 2003) and a questionnaire developed to measure social expectations towards robots, which is shown in Table 1 along with its dimensions. This latter questionnaire is used to investigate social expectations of robots along three dimensions. The first dimension, *Equality*, is comprised of the items “Friend” and “Colleague”. High scores on this would suggest that the participant would expect the robot to act in manner suggesting an equal (social) footing to themselves within interactions, whilst a low score would suggest the opposite (i.e. that the robot adopts a more deferential role). The second dimension, *Control*, is comprised of the items ‘Servant’ and ‘Tool’. High scores along this dimension suggest that the user would expect that the interaction with the robot is one in which the user will exert a high degree of control, while a lower score would suggest that the robot is expected to act in a more autonomous manner. The third dimension measured using the questionnaire is that of the *Pet* dimension which suggest an expectation that the robot should be like a pet within the interaction. Please refer to (Koay et al. 2014) for a more detailed discussion of these issues.

A second version of this questionnaire was given to participants after the interaction, in order to examine their social perception of the robot.

How would you like to interact with robots in the future? *I would like to interact with them as a…?*	Pre-test question
What was the robot like in this study? *It was like interacting with a….?*	Post-test question
Item Friend	Dimension Equality
Servant	Control
Pet	Pet like
Colleague	Equality
Tool	Control

Table 1: Social Expectations Questionnaire. The questions shown at the top of the table were asked prior to, and after the experimental study, respectively. Possible choices of answers are shown in the bottom part of the table.

The experimenter then took over and the psychologist retired to a different room. The experimenter explained the purpose of the experiment, the nature of the sensorised house and the capabilities of the robot (in this experiment the robot capabilities were restricted to moving to differing locations and speaking, although the tray and arm/gripper were visible).

The robot had previously been taught to approach the experimenter and participant and to introduce itself by saying "welcome to the robot house". This gave the experimenter a chance to explain the robot capabilities and for the participant to see the robot in action for the first time.

Examples of three sets of behaviours, each with increasing complexity, were shown to participants (the behaviours are shown in Table 2, Table 3 and Table 4). The behaviour relating to "answering the doorbell" in Table 2 was used by the experimenter to show the participant how to use the teaching GUI.

The tables below show three sets of behaviours presented to participants. Higher numbered sets are more complex to teach. Each participant chose one behaviour from each set in order to teach the robot.

Robot Behaviour	Achieving
Whenever you open the microwave oven, make the robot come to the kitchen and wait outside.	*Assistance – physical*
If the TV is on, and you are sitting on the sofa, make the robot join you at the sofa.	*Robot as Companion*
If the doorbell rings and the TV is on, make the robot say "There is someone at the door" and then go to the hallway.	*Assistance – medical - hearing impairment*

Table 2: Taught Behaviours Set 1. The behaviours above comprise of two taught components – 1) react to a house sensor event and 2) make the robot carry out some physical action.

Robot Behaviour	Achieving
Make the robot come to the table and remind you to always call your friend "Mary" on Thursday at 2pm.	*Re-ablement - avoid social isolation*
On Mondays, Wednesdays and Fridays make the robot stand in the hall and remind you that your lunch is being delivered at 12:30pm.	*Assistance – medical – memory*
If you are sitting on the sofa for more than 30 minutes, make the robot come to the sofa and tell you to take some exercise. Make the robot do this again after another 30m if you are still sitting there.	*Re-ablement - avoid sleeping too much during the day*
Make the robot come to the table and remind you to take your medicine at 12pm and 4pm every day, yellow pills at 12, pink pills at 4pm.	*Assistance – medical - medicine*

Table 3: Taught Behaviours Set 2. The behaviours for set 2 above are more complex than those in set 1 as they additionally comprise diary and temporal functions associated with both physical and verbal robot actions.

Robot Behaviour	Achieving
Make the robot come to the sofa and tell you to "move about a bit", if, in the afternoon , you have sat on the sofa for longer than 1 hour continuously	*Re-ablement - avoid sleeping too much during the day*
If it is after 9pm, and you have left the sofa for over 5 minutes and the TV in on, make the robot go to the hall entrance and say "turn off the TV".	*Assistance – medical – memory*
If the microwave has been open or on for more than 1 minute, make the robot come to the table and tell you that the microwave is open or on. Make the robot remind you every minute until the microwave is turned off and door is closed.	*Assistance – safety*

Table 4: Taught Behaviours Set 3. The behaviours for set 3 above are more complex than those in set 2 as the diary and temporal functions are now either open-ended (e.g. "Afternoon") or require repetition together with either physical or verbal robot actions.

During the teaching process the experimenter stayed with the participant and helped them when asked. Given that none of the participants had ever interacted with a robot before, and that the teaching GUI was entirely new to them, we felt that this was a necessary requirement. Furthermore, part of the post experimental questionnaire asked them to indicate whether they thought they could continue to use the teaching system without the help of the experimenter. The participant's use of the teaching system was also videoed for later analysis.

Having taught the robot the new behaviour the experimenter then invited the participant to test it. If the behaviour operated successfully then the participant moved on to teaching the next behaviour in the subsequent set. Alternatively they could modify the existing behaviour and re-test.

Having taught all three behaviours (one from each set) the experimenter retired to another room and the psychologist returned and asked the participant to fill in a post evaluation questionnaire based on Brooke's usability scale (Brooke 1996) (see Table 5). We felt that this separation of duties between the experimenter and psychologist was necessary to avoid putting any pressure on the participant when they were completing the evaluation questionnaire.

I think that I would like to use the robot teaching system like this often.
I found using the robot teaching system too complex.
I thought the robot teaching system was easy to use.
I thought the robot teaching system was easy to use
I think that I would need the support of a technical person who is always nearby to be able to use this robot teaching system.
I found the various functions in the robot teaching system were well integrated.
I thought there was too much inconsistency in the robot teaching system.
I would imagine that most people would very quickly learn to use the robot teaching system.
I found the robot teaching system very cumbersome to use.
I felt very confident using the robot teaching system.
I needed to learn a lot of things before I could get going with the robot teaching system .

Table 5: Modified Brooke's Usability Scale. The table shows the questions posed to participants specifically relating to usability. All answers were based on a 5-point Likert scale.

A subsequent 'ad-hoc' questionnaire (see section4.2.3) was also completed which focused specifically on the usefulness of the robot and teaching system.

After completion of the questionnaire the participant was invited to ask questions if they wished about the experience, the house, the robot etc. In fact, all of the participants were very interested to see how the house and robot worked.

4.2 Results

4.2.1 Demographics

There were 20 participants in the study, 16 female and 4 male. The mean age in the sample was 44 years (SE=15.3), with a median age of 49 years. The computer usage of the participants can be found in Table 6: Computer Usage in the Sample, which suggests that the majority of participants used computers for work/studies as well as for social reasons.

Activity	**Yes**	**No**
Work or Study	18	2
Socialising	19	1
Recreation	8	12
Programming	0	20

Table 6: Computer Usage in the Sample

There was a split in the sample however, in that about half of the participants used computers for recreational reasons, such as games. None of the participants programmed computers. The mean number of hours spent on computers in the sample was 35 hours (SE=2.98) with a median number of hours of 33. Only one of the participants had had any experience with robots of any sort. Table 7 shows the responses to the TIPI in the sample. Table 8 shows the responses to the Social Role questionnaires which indicate that participants initial expectations of the social roles of robots did not differ significantly from their perception of the robot within the actual interaction.

	Mean	**SD**
Extraversion	4.38	1.48
Agreeable	5.35	1.14
Conscientious	5.83	1.15
Emotional Stability	4.85	1.36
Openness	5.17	1.10

Table 7: Personality in the Sample

Role	Stage	Mean	SE	Diff	95% CI	t(19)
Equal	*Expect* *Actual*	3.1 2.88	0.21 0.24	0.22	-0.35 to 0.8	0.82
Control	*Expect* *Actual*	3.7 3.52	0.17 0.18	0.18	-0.29 to 0.64	0.79
Pet	*Expect* *Actual*	2.35 2.21	0.25 0.22	0.21	-0.54 to 0.96	0.59

Table 8: Initial Social Expectations and Actual Social Perception

4.2.1.1 *Responses to the System Usability Scale (SUS)*

The mean participant response to the System Usability Scale regarding the teaching interface was 79.75 (SE=2.29), and the median response was 76.25. These scores were significantly higher than the "neutral score" of 68 ($t(19)=5.12, p<.01$).

A multiple regression analysis was conducted in order to investigate demographic predictors of SUS responses to this task. After removing non-significant predictors, the final model had an adjusted r^2 of .28, and predicted SUS scores significantly ($F(2,17)=4.70, p<.05$). The model is described in Table 9 and suggests that both age and scores on the *Conscientiousness* personality trait were associated with lower scores on the SUS for this task.

Predictor	β	SE	T(19)	p
Intercept	0.00	0.00	0.00	1.00
Age	-.49	.20	-2.48	< .05
Conscientiousness	-.40	.20	-2.23	< .05

Table 9: Predictors of System Usability Scores

4.2.2 Responses to Ad-Hoc Questions

Participant responses to the ad-hoc Likert items can be found in section 4.2.3. All participant responded '*Very Useful*' or '*Useful*' when asked if they thought it useful to teach a robot. In addition all participants answered '*Definitely Yes*' or '*Yes*' when asked if they thought that *they* would be able to teach the robot, if *they* would do so for a relative, and that *they* would find it useful to customise the tasks of a robot beyond a set of standard tasks. The participants did not, however, agree as strongly on whether or not the robot should be completely set up by someone else, with a wider range of responses from the participants.

Participants also responded that they were overall '*Very Comfortable*' or '*Comfortable*' with a robot informing them that there was a problem in their house, and 17 out of the 20 participants answered that they were at least '*Comfortable*' with the robot informing a third party about an unresolved problem, but there was less agreement regarding having a robot suggest that they play a game or exercise.

As these were ordinal Likert-items, linear regression analyses were not performed as for the SUS scores. Instead a series of exploratory Spearman's correlations were performed to investigate relationships between the items that there were disagreements regarding in the sample, and the measures described in the demographics section.

For wanting the robot already set up, there was a significant correlation between this and the '*Equality*' dimension of the initial social expectations (*rho*(20)=.70,*p*<.01), and a correlation approaching significant between this and the '*Emotional Stability*' personality trait (*rho*(20)=.40,*p*=.08). Both of these correlations indicated that participants with higher scores along these dimension were less likely to want the robot fully set up by someone else. There was also a trend approaching significance for this item and '*Age*' (*rho*(20)=-.37,*p*=.10$), in which older participants were more likely to want the robot already set-up.

There were no significant relationships between 'comfort' with the robot suggesting that 'one take more exercise' and the demographic measures, but there was a trend approaching significance for the *Control* dimension of the Actual Social Role Perception of the robot (*rho*(20)=.40,*p*=.07). This trend suggested that participants who rated their interaction with the robot highly along this dimension were less comfortable about the robot making such suggestions.

There was a significant relationship between the *Equality* dimension of the initial social expectations and 'Comfort' with a robot suggesting that 'one play a game with it' (*rho*(20)=-.69,p<.01), suggesting that participants that scored highly on this dimension were more comfortable with the robot making such suggestions.

There was a significant relationship between *Age* and 'Comfort' regarding the robot contacting a third party in case of a problem (*rho*(20)=-.53,*p*<.05), where older participants were more comfortable with this.

4.2.3 Frequencies of responses to the ad-hoc Likert Items

Do you think it is useful teach a robot?

Very Useful	*Useful*	*Neither*	*Not Useful*	*Not at all*	*Median*
18	2	0	0	0	1.00

Do you think that you would be able to teach the robot?

Def. Yes	*Yes*	*Neither*	*No*	*Def. No*	*Median*
10	10	0	0	0	1.50

Would you be willing to teach the robot for someone else, e.g. if you were a relative or carer of the other person?

Def. Yes	*Yes*	*Neither*	*No*	*Def. No*	*Median*
14	6	0	0	0	1.00

Do you think that robot should already have been completely setup by someone else?

Def. Yes	*Yes*	*Neither*	*No*	*Def. No*	*Median*
1	3	4	11	1	4.00

Do you think that the robot should be able to carry out standard tasks but it would be useful to be able to customize it?

Def. Yes	*Yes*	*Neither*	*No*	*Def. No*	*Median*
13	7	0	0	0	1.00

Is it useful to know what the robot can already do?

Def. Yes	*Yes*	*Neither*	*No*	*Def. No*	*Median*

12	8	0	0	0	1.00

How would you feel about having a robot suggesting that you take more exercise?

V. Comf.	*Comf.*	*Neutral*	*Uncomf.*	*V. Uncomf*	*Median*
9	8	2	1	0	2.00

How would you feel about having a robot suggesting that you play a game together e.g. a video game or chess/draughts?

V. Comf.	*Comf.*	*Neutral*	*Uncomf.*	*V. Uncomf*	*Median*
6	11	2	1	0	2.00

How would you feel about having a robot warning you that there was a problem in the house e.g. fridge left open or hot/cold taps running or TV left on?

V. Comf.	*Comf.*	*Neutral*	*Uncomf.*	*V. Uncomf*	*Median*
18	2	0	0	0	1.00

How would you feel about having a robot informing someone else that there was a problem in the house e.g. by texting them, if the problem had not been resolved?

V. Comf.	*Comf.*	*Neutral*	*Uncomf.*	*V. Uncomf*	*Median*
12	5	2	1	0	1.00

4.3 Robot Teaching Formative Study - Discussion

4.3.1 Summary of Results

The results from the SUS suggest that participants found the interface easy to use. Moreover, all participants indicated that they felt able to use a system like this to teach the robot, and willing to use such a system to set-up behaviours for an elderly relative or person in their care. These are encouraging results which suggest that further development of the robot teaching system is warranted.

In terms of individual differences, there are some salient relationships. The relationship between Age and SUS scores are not unexpected. The older members of the sample found the system more difficult to use than the younger participants. Related to this is the impact of age on the ad-hoc item regarding wanting the robot to be already set up by someone else. Here, older participants were more likely to want the robot being fully set-up than younger participants.

Taken together, this suggest that the current stage of this teaching system may be better suited for use by carers and relatives of elderly people to set up the robot's behaviours for them, but that it needs to be further developed in order to be more suitable for the use of elderly people themselves.

The relationship between items covering the possibility of the robot contacting third-parties in case of problems, and *Age* is also interesting (and we envisaged that this would be a key item that may be taught to the robot). While one explanation for this result may be that older participants were closer to having to consider these scenarios in their own lives than their younger counterparts, a more likely explanation may be that the older portion of the sample were more likely to have had more experiences with caring for elderly parents or other relatives and so might identify more strongly with the third party that is to be contacted. Some of the informal responses from participants during the debrief of the study did reference such experiences.

5 Formative Evaluation of the Visualisation System

5.1 Brief Recap the Behaviour Visualisation System

Equipping a robotic companion with a visualization tool for episodic memory is an opportunity to have a robot provide memory prosthesis. Such memory visualization can support the user in remembering past events from the human-robot interaction history. Potentially, this ability to explore interaction histories could enable elderly persons as well as third parties (e.g. technicians, carers, family and friends) to monitor, maintain and improve the robot's abilities and services.

The robot's interaction history is stored in, and retrieved from, the robot's episodic memory, and visualisation of user past experiences can be achieved through an embedded touch interface. Pictures of the environment are captured in a particular moment during the robot's goal execution and linked to its episodic memory, which is used to create interaction histories for the purposes of event visualization.

The structure of the action history consists of the environmental sensory state which becomes part of the visualization history. Memory visualization is fully integrated into the complete sensory and action architecture and can be generated on the basis of both behavioural execution or on a timed basis. When an action sequence (that part of the behaviour which runs on the robot) is executed, a record is added to the history log table with the behaviour name, the location of the robot, and the current time stamp. The collection of other relevant data is then started, i.e. an image from the robot's on board camera is taken as well as a snapshot of the state of the house sensor network. When the user opens the memory visualization interface, the system retrieves a chronological list of all actions that have been recorded. When a user selects an action, they are presented with the previously mentioned recorded information. See the Figure 2, Figure 3 and Figure 4 for examples of the visualisation system in action.

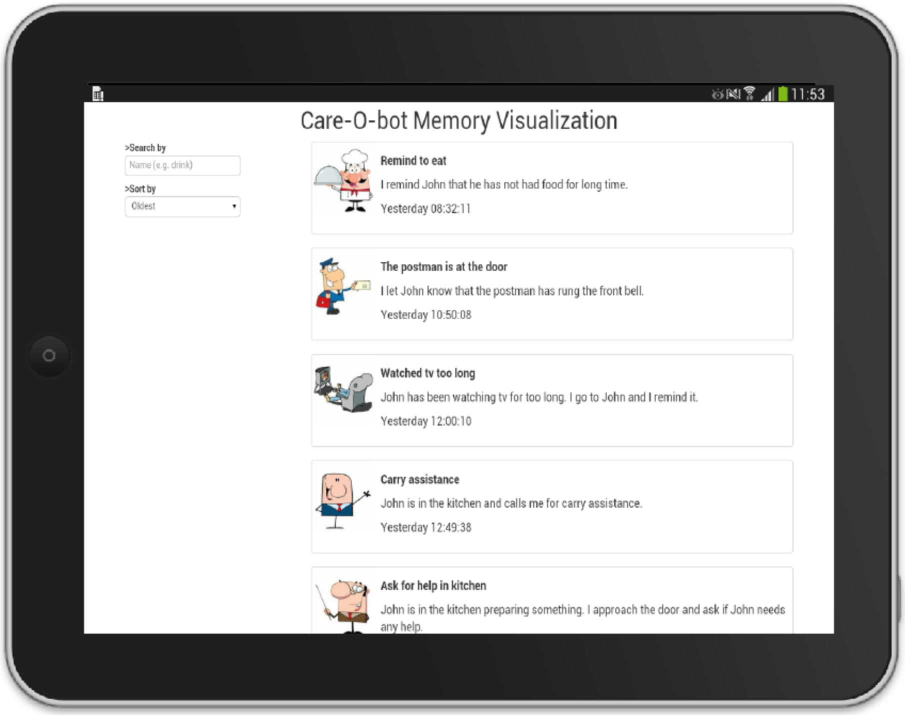


Figure 2: Initial screen

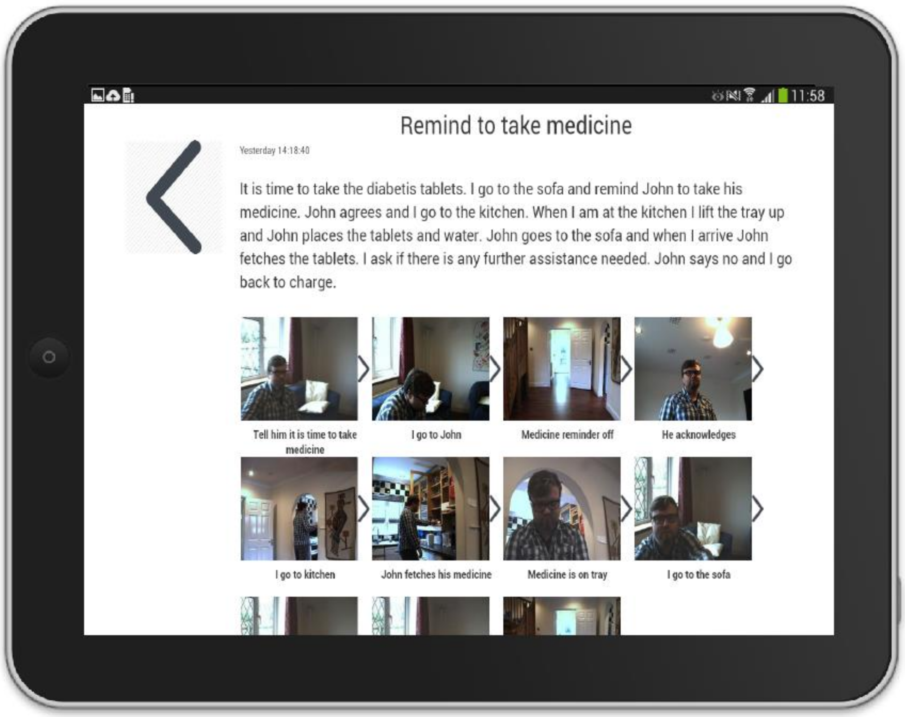


Figure 3: Thumbnail photos of the chosen activity

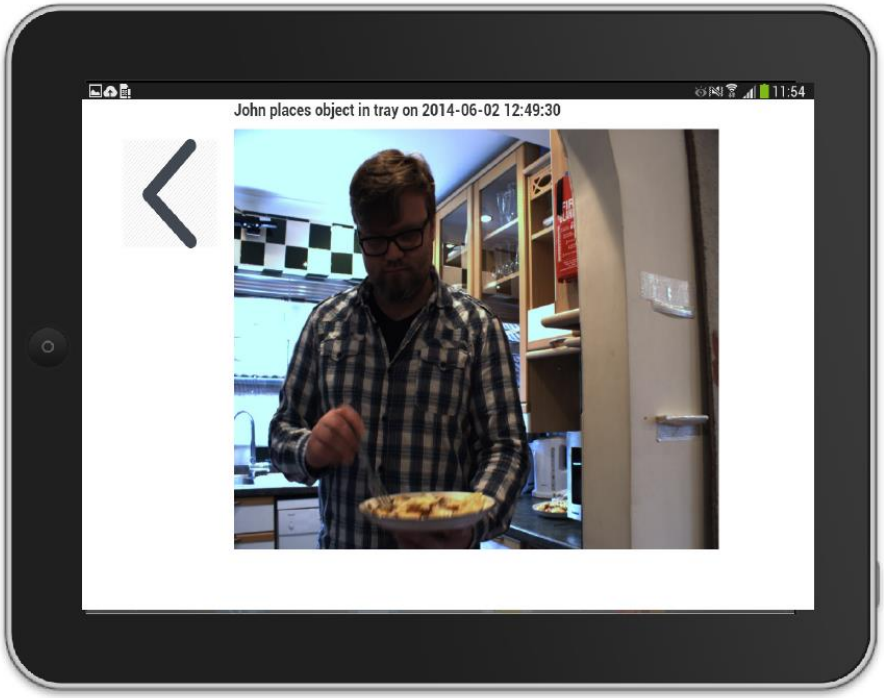


Figure 4: Zoom in on a particular part of the activity

5.2 Procedure

The procedure for the formative evaluation of the Memory Visualisation System was carried out in conjunction with the formative evaluation of the Robot Teaching System as explained in Section 4. After the participant got used with the robot and the Robot Teaching system they were given a set of instructions on how to use the Robot Memory visualisation system (see Figure 5). Once familiar with the system they were then asked to answer a series of questions (see Figure 6) relating to one day's activity in the house, where the robot had interacted with a person called "John".

Instructions
The robot has been interacting previously with John. You can look at the robot's memory what John has been doing at the Robot House using the visualization tool. The experimenter will first show you how to use this simple interface with an example.
Have a look at what John has been doing with the robot.
With the memory visualization tool we will ask you to fill in a questionnaire and then one of the experimenters will ask you a few questions.
Please remember that we are testing the robot and the visualization system, we are NOT testing you!

Figure 5: User instructions

The first questionnaire was a set of questions which the participants could only answer using the memory visualization system:

Question
What time did the postman arrive?
What colour was the parcel the postman delivered?
Where did the robot go after helping John with the parcel?
What was John doing when the robot reminded him to drink?
What colour was the cup John drank from?
What was John eating when he asked for assistance?
What time did John have lunch?
Where does John keep his medicine?
Did John call his daughter Mary?
What was John doing in the kitchen when the robot asked if he needed help?

What did John do when the robot reminded him that he had watched TV for too long?

Figure 6: Questions on one day's activity in the house.

After the interaction with the Memory Visualisation System and answering the questions the participants were given a System Usability questionnaire to evaluate the usability and the features of the Visualisation System (see Figure 7).

Statement
I think that I would like to use the memory tool like this often.
I found using the memory tool in this session too complex.
I thought the memory tool was easy to use in this session.
I think that I would need the support of a technical person who is always nearby to be able to use this memory tool like I did in this session.
I found the various functions in the memory tool for this session were well integrated.
I thought there was too much inconsistency in the memory tool for this session.
I would imagine that most people would learn to use the memory tool as it was used for this session very quickly.
I found the memory tool very cumbersome to use in this session.
I felt very confident using the robot in this session.
I needed to learn a lot of things before I could get going with the memory tool as I did in this session.

Figure 7: System Usability Scale for the Visualisation evaluation

Afterwards a set of questions shown in 5.3.1 were provided in order to test the usefulness of the Robot Memory Visualisation System, with the answers in different Likert scales (very useful – useful – neither useful nor useless – not useful at all).

5.3 Results

The System Usability Scaly (SUS) scores for the Memory Visualisation interface ranged from 37.5 to 95. The mean score was 74.85 and the median score was 75. This was significantly different from the expected average of 68 ($t(19)= 2.79$ $p< .05$).

5.3.1 Add-hoc Likert Scales

The following ad-hoc questions were also given to the participants:

Interface Specific Questions

The responses to the interface specific questions can be found below:

Question			Frequencies			Mean	SD	Median
	Very Easy	Easy	Neutral	Difficult	Very Difficult			
Ease of Use	11	6	0	0	0	1.35	0.49	1
	Very Clear	Clear	Neutral	Unclear	Very Unclear			
How Clear was the Interface	4	9	2	2	0	2.12	0.93	2
	Definitely Yes	Yes	Don't know	No	Definitely No			
Would like video	4	3	3	7	0	2.76	1.25	3

The responses to the questions regarding the utility of such a system are shown below:

Question			Frequencies			Mean	SD	Median
	Very Useful	Useful	Neither useful nor useless	Not Useful	Not Useful at all			
Reviewing interactions useful	13	3	1	0	0	1.29	0.59	1
	Very Useful	Useful	Neither useful nor useless	Not Useful	Not Useful at all			
Useful for recalling	10	7	0	0	0	1.41	0.51	1

interactions								
	Very Useful	Useful	Neither useful nor useless	Not Useful	Not Useful at all			
Useful for finding robot errors	8	9	0	0	0	1.53	0.51	2
	Very Useful	Useful	Neither useful nor useless	Not Useful	Not Useful at all			
Useful for starting conversations	2	11	3	1	0	2.18	0.73	2
	Definitely Yes	Yes	Don't know	No	Definitely No			
Will give better overview of daily routines	7	8	0	2	0	1.82	0.95	2
	Definitely Yes	Yes	Don't know	No	Definitely No			
Will help you remember routines	8	6	1	2	0	1.82	1.01	2

Security and Comfort Questions.

The responses to the security and comfort questions can be found below.

Question	Frequencies					Mean	SD	Median
	Very Comf.	Comfortable	Neutral	Uncomf.	Very Uncomf.			
Feel about robot storing information	4	9	3	1	0	2.06	0.83	2
	Very Comf.	Comfortable	Neutral	Uncomf.	Very Uncomf.			
Feel about being reminded	8	8	1	0	0	1.59	0.62	2
	Definitely Yes	Yes	Don't know	No	Def. No			
Would monitor family members?	12	5	0	0	0	1.29	0.47	1
	Definitely Yes	Yes	Don't know	No	Def. No			
Robot to store conversations	2	8	5	1	1	2.47	1.01	2

Responses to the open-ended questions can be found below:

IS THERE ANYTHING YOU WOULD CHANGE?
No
static pictures could be enlargable to more detail
Clarity of pictures
Not particularly
None
Improve pictures - better quality
No
Enlarge the pictures
Could be helpful for family member to see videos of the events instead of pictures
Having a zoom option on the photographs would be useful for finer details/clarity
Colour of images could be clearer
The video footage would be helpful in a situation where a person wants to check how an elderly person has been throughout the day as the could replay the day
Only that I couldn't see the photos as clearly as thumbnails so would want to view them larger.
No being able to make tea would be brilliant though
WHAT OTHER INFORMATION WOULD YOU ADD TO THE EVENTS, IF ANY?
Something about how well the robot is reacting and working
None
I would add timestamps next to the pictures
The picture showed him going to the fridge for breakfast but not actually eating it. My mother in law used to put her food in the toilet
When somebody approaches the front door could you have the CCTV automatically switch on and related to the robot even if they don't ring the doorbell so the carer can see who has been to the property
no sure if it would record how many times he had to ask them to 'take medicine' for example that could be

useful.
Waking/getting up time or breakfast and getting to bed time. Any exercise/movement activity e.g. went for a 30 minute walk - any problems encountered by "John" e.g. could not reach the cupboard
lock door at night
It would be good if this footage were available remotely i.e. if a person is at work some distance from the elderly person they could view footage of how the elderly person is e.g. did they get out of bed/ go back to bed; have breakfast have a wash; get their tablets - worse still if they have taken a fall or let a salesman into the house who will avail of their vulnerable mindset & sign them up for something. This would be useful security for the elderly person and peace of mind for the carer during the hours when the carer is not physically present.
Video sounds good but would take up lots of memory space so if this was an issue I would settle for still photos only
Spilling a drink; hearing your phone ring and going to a spot near it; washing machine finished find out football score and tell me

5.4 Visualisation System Formative Study - Discussion

All participants found the interface easy to use, and the majority of participants found it clear, with a small minority stating that it was unclear. The sample was divided as to whether or not they wanted video in addition to images.

The responses regarding the utility of the memory visualisation system suggested that the participants were overall positive to all aspects of the utility of the memory visualisation system. Two participants did not think that it would aid in getting a better overview of daily routines nor that they would aid in remembering them.

The majority of participants were comfortable with the robot storing information about them and about being reminded about events. In addition all participants stated that they would use the robot to monitor family members. The sample was more divided as to whether or not they would want conversations stored.

There was a variety of suggestions for what else the robot could store and report back to the user, timestamps, the internal states of the robot, specific problems encountered during the day as well as registering visitors to the property. One participant wanted the footage to be available remotely, so that family members could review it off-site, but one comment questioned whether or not the primary user might be able to trick the system so that it would seem that they performed certain tasks.

In the open-ended questions the most common type of comments were regarding the size, and clarity of the images. The participants wanted the photos to be larger and/or being able to zoom in on parts of the images.

6 Developments to Support learning

In section 3.3 above, we described the evaluation of *teaching* component of *co-learning*. In this section we will outline the *learning* part. We approach this problem by considering how the robot can learn through its experiences in the house. Both of these areas were described and supported with a literature review presented in deliverable D3.1 and D3.2.

6.1 Learning From Experiences in the House

In this work the robot learns through its experiences in the house and is based on work by Saunders et al., (2007) which was previously outlined in Deliverable 3.2. The approach is contingent on the house occupant indicating to the robot that activities are underway. For example, the person might indicate that they are now "preparing food".

Activities typically have a nested and sometimes hierarchical nature. For example, "preparing food" might also include making a hot or cold drink or using the toaster. An example of these nested tasks is shown in Figure 8.

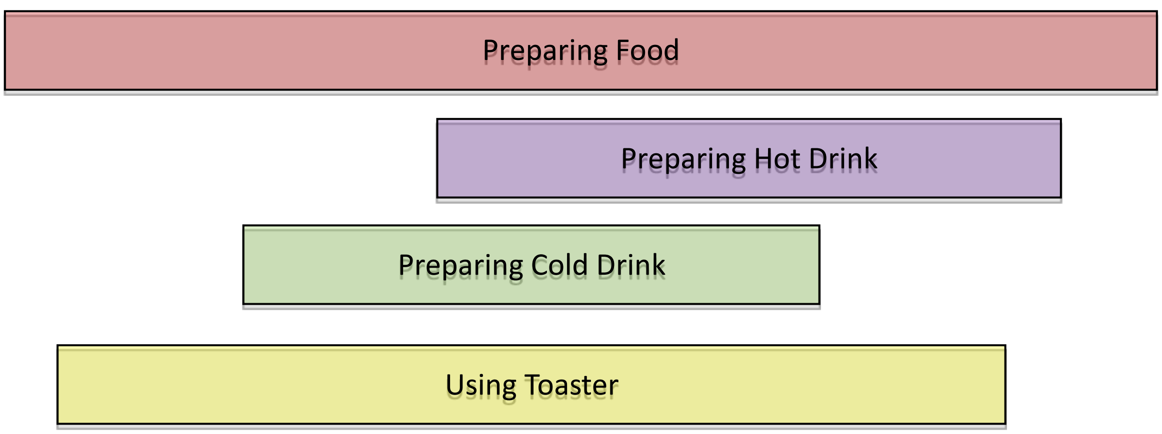


Figure 8: Nested activities under the banner of 'preparing food'. Time is moving from left to right.

The start and end times and durations of the main task and the sub-tasks are completely variable. However, when any of the sub-tasks are active (e.g. using toaster) the main task must also be active (i.e. preparing food).

Consider that the person has indicated to the robot that they are "preparing food" and at some point they also indicated that they are now "using the toaster". If the robot learns the set of sensory activities associated with these tasks it should be able to recognize them when they occur again in the future. Thus the robot would recognize when the toaster is active and infer not only that "using the toaster" is true but also that "preparing food" is true.

Given that these activities can be recognized by the robot, it would then be possible to exploit these in the teaching system outlined in Section 3.3 and the person would now be able to teach the robot based on the higher level semantics associated with the task. For example, the user might teach "When I am *'Preparing food'*, the robot should come to the kitchen and raise its tray".

The relationship between learning and teaching is shown in Figure 9 where the learning system provides symbolic entries into the set of existing robot competencies. These can then be exploited by the teacher to create new behaviors on the robot.

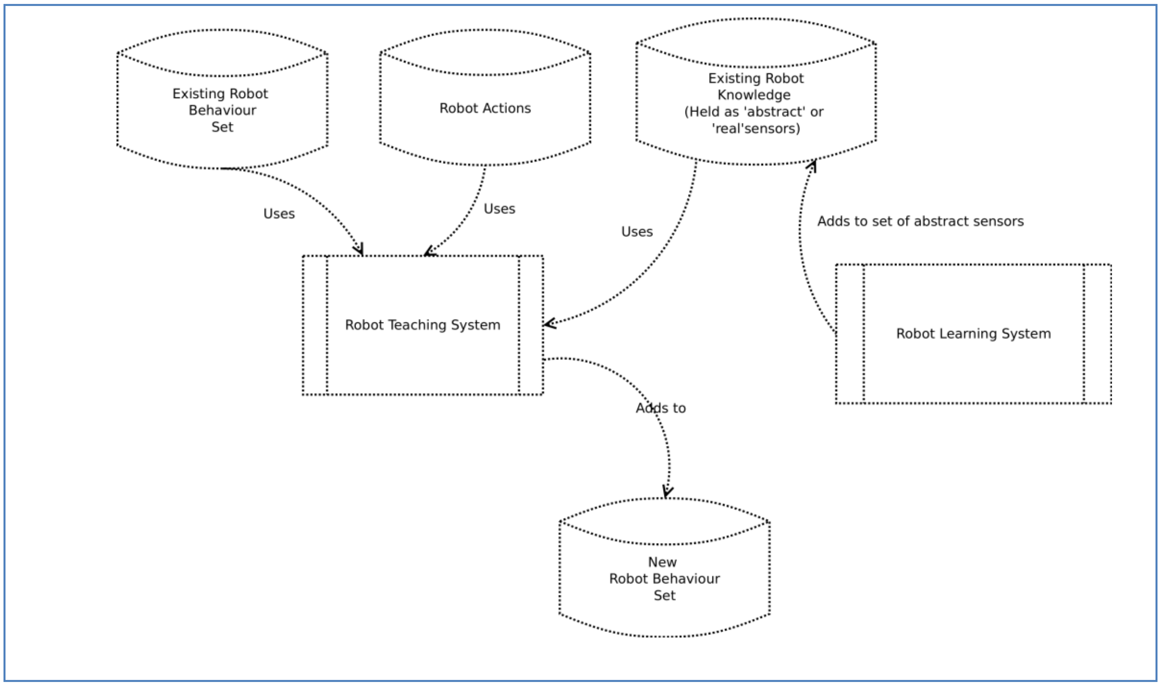


Figure 9: Relationship between Learning and Teaching

The challenges for a learning system are therefore to be able to recognise that learnt situations can be active in parallel, have an implicit nested hierarchy and that higher levels in the hierarchy (typically) represent higher level of semantic knowledge. These need to be represented as lexical symbols in the memory architecture which the teacher can then exploit.

6.1.1 Approach to Learning

In order to learn typical activities in the house the robot needs to recognize when these situations re-occur. This recognition would be primarily based on the current sensory state of the house, however, in more complex circumstances both the historical sensory state and a predicted future sensory state may also be necessary (for example, in historical terms, to recognize that the postman called this afternoon, or in the predicted sense, that the house occupant is likely soon to go to bed). In the work presented in this deliverable we consider the only the current sensory state however work on a predicted sensory state (for example using sequential data mining algorithms) is suggested for future studies and preliminary work on this is shown in the Appendix.

We also have to consider that certainty of situations cannot always be represented by a simple true/false dichotomy. For example, if I am in the kitchen it is likely I am preparing food, but it is not a certainty. The confidence of the task assessment by the robot has to be considered.

6.1.2 Related Work

A discussion of related work on learning was presented in D3.2. In this document specific machine learning approaches to recognition of smart home activities will be discussed below and how our approach exploits and extends some of the approaches to allow house participants to personalize the robot for their own needs. A recent detailed survey of a wide range of technologies supporting ambient assisted living can be found in (Rashidi & Mihailidis 2013).

Our approach falls under the banner of *Ambient Activity Recognition* in that house resident activities are modelled by analyzing a sequence of ambient sensor events. The various approaches to this research area typically apply supervised machine learning techniques such as decision trees/rule induction (Chen et al. 2010) (as is used in the studies presented in this document), HMM's and dynamic Bayesian networks (Lester et al. 2005), template matching techniques such as *k*-NN (Saunders et al. 2006) or dynamic windowing techniques (Krishnan & Cook 2014) . Sensor data is typically pre-labelled by an external observer (although some techniques also search for common patterns in daily activities (Gu et al. 2009). Our approach differs from a strict supervised learning approach in that the house resident is responsible for 'labelling' the data and does this by simply carrying out the activity whilst the system records and automatically labels the sensory data accordingly. Furthermore, the newly acquired activity can be subsequently used for direct robot teaching. Activity recognition is based on streaming vectorised sensor data – an approach which allows multiple activity patterns to be recognized in parallel.

6.1.3 Approach

The current memory system as a whole is based on rule sets. As outlined in previous deliverables, these are human readable and in many cases taught by the human using the teaching system. Ideally a learning system should also be human readable to allow query by the user. We therefore decided to employ a rule induction approach to learning based on Quinlan's C4.5 Rule induction algorithm (Quinlan 1993). In fact we use the latest version C5.0.

6.1.4 Testing the plausibility of the approach with existing data sets

In order to verify the plausibility of our approach we exploited some existing end user behavior data available from previous studies in the UH robot house (Duque et al. 2013). In these previous studies 14 participants were asked to carry out a series of typical daily activities within the house. Each participant took part in two sessions of approximately 45min duration each.

In the first (training) session the experimenter suggested to the participant particular tasks that should be carried out. In the second (test) session the experimenter asked the participant to carry out any of the tasks (or none) at their own discretion.

During the experiment all house sensory data was recorded and all sessions videotaped. Subsequently the video tapes were annotated by both Duque and an external observer (with subsequent appropriate inter-observer correlation carried out) and marked with the task and sub-task start and end times. These were then matched against the recorded sensory data.

Duque's aim in these experiments was to *manually* create a rule-set, derived from the training sessions, which could be applied to the test data and accurately predict the activity that was being carried out by the participant. This rule set was constructed and applied to the test data resulting in recognition accuracy (based on precision, recall and accuracy) of over 80%. A diagram of this work is shown in Figure 10.

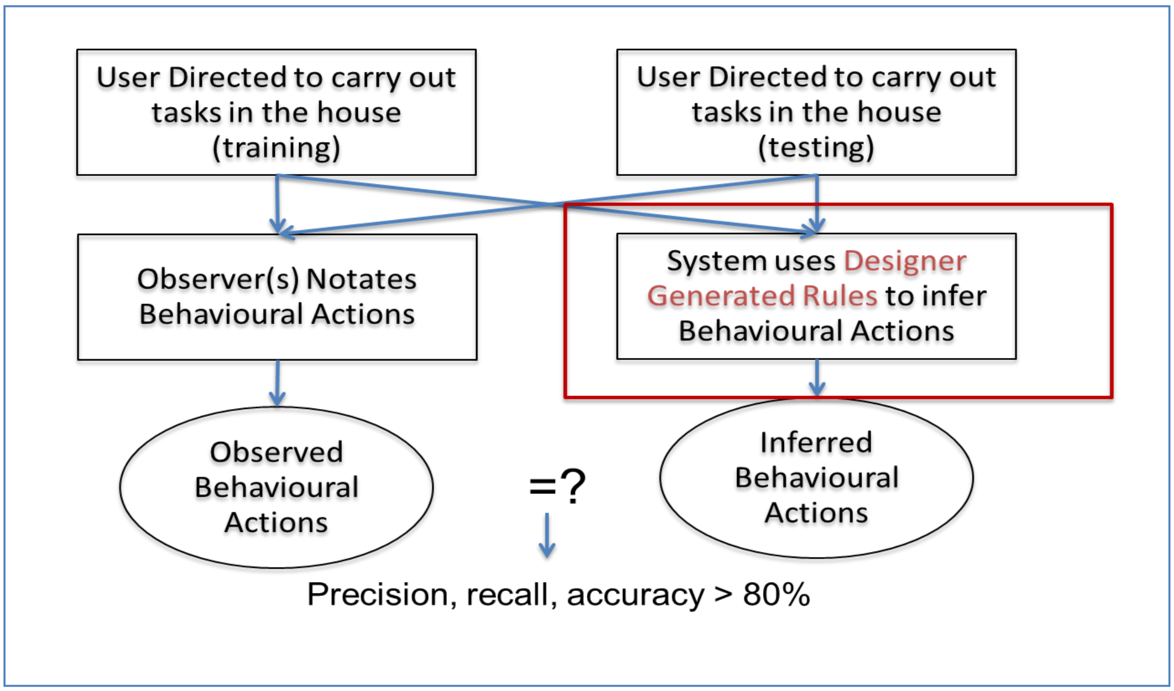


Figure 10: Duque's previous work on deriving rule sets.

In the work presented in this deliverable we tested the plausibility of the approach by replacing the *designer generated rules* with rules *automatically derived* using the C5.0 algorithm. We then assessed the performance of this approach.

6.1.5 Existing data sets

An example for the first four tasks for the first four participants for both the test and training sets is show in Figure 11.

For each participant (only 4 participants are shown), the training tasks are shown at the top of each block and the test tasks shown at the bottom. The continuous lines show the length of the activity e.g. for user001 'preparing food' in the training scenario starts at 0.15 and ends at 5.10. Each colour represents the task – red = preparing food, purple = preparing hot drink, green = preparing cold drink and yellow = using toaster.

Note that in some cases (for example user003) the participants carried out a task in the training scenario, but chose a different task in the test scenario. This implied that the learning system would have no data available to deduce the activity (as it has never seen it before).

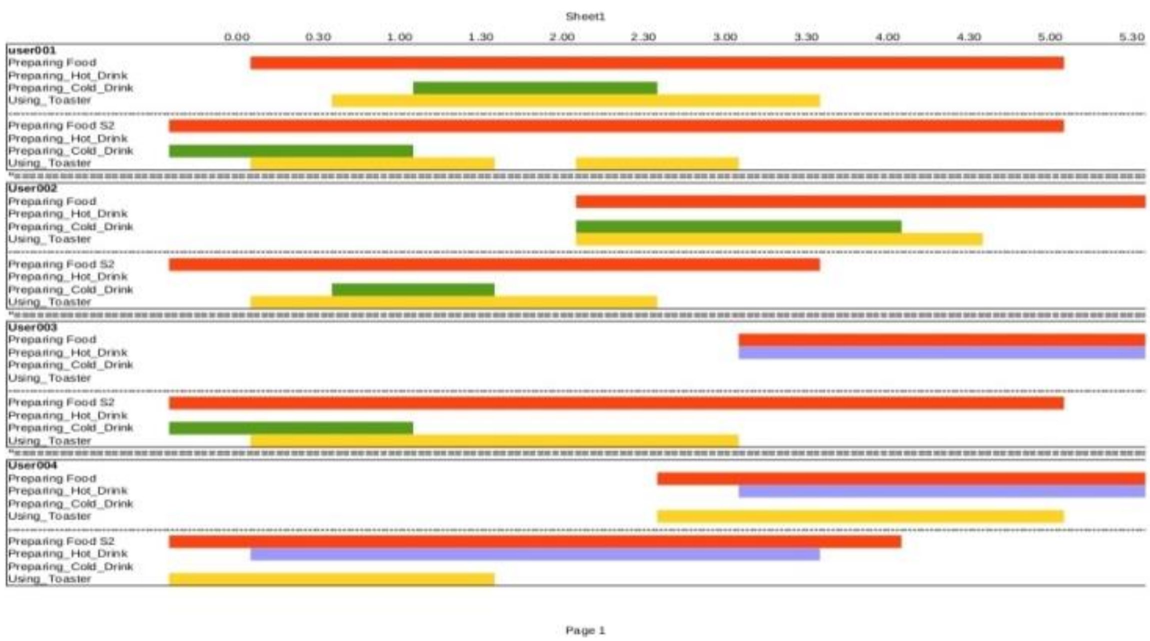


Figure 11. First four participants in Duque's data set. For each participant the top section shows the training data and the bottom section shows the test data. Only the first four tasks are shown.

6.1.6 Results from Existing Data Sets

The training data for the 14 participants was used to train a learner using C5.0 with boosting over 10 trials. The learner was then applied to the test data and the resulting performance analysed. A typical derived rule set (in this example for 'preparing a hot drink') is shown in Figure 12. The rules are interpreted as follows (see left box, trial 0):

If kettle **is On**
Then preparing Hot Drink
Else
If water pipe sink cold **is on**
Then preparing Hot Drink
Else
If toaster **is on**
Then preparing Hot Drink
Else undefined

'PreparingHotDrink' Trial 0

10 sets of rules (trials) are generated from the data set and statistics of performance evaluated as to the efficiency of the rules in each trial (see Figure 13 for the boosting and confusion matrix statistics). When a new case is to be classified, each trial votes for its predicted class and the votes are counted to determine the final class.

```
C5.0 [Release 2.07 GPL Edition]        Thu Jul  3 12:01:16
2014
---------------------------

    Options:
        Application `Inferences'
        Boosted classifiers

Class specified by attribute `activity'

Read 1167 cases (65 attributes) from Inferences.data

----- Trial 0:  -----

Decision tree:

Kettle = on: Preparing_Hot_Drink (141)
Kettle = off:
:...Water Pipe Sink Cold = on: Preparing_Hot_Drink (17)
    Water Pipe Sink Cold = off:
    :...Toaster = on: Preparing_Hot_Drink (46/15)
        Toaster = off: Undefined (963/19)

----- Trial 1:  -----

Decision tree:

Sofa Seatplace 0 = on: Undefined (558.3)
Sofa Seatplace 0 = off:
:...location = Living Room(2): Undefined (31.8)
    location = Kitchen(10): Preparing_Hot_Drink (492.1/176.4)
    location = Dining Room(25): Undefined (84.8)
```

```
----- Trial 2:  -----

Decision tree:

Kettle = on: Preparing_Hot_Drink (84.9)
Kettle = off: Undefined (1082.1/166)

----- Trial 3:  -----

Decision tree:

Sofa Seatplace 0 = on: Undefined (351.1)
Sofa Seatplace 0 = off:
:...Kettle = on: Preparing_Hot_Drink (67.2)
    Kettle = off:
    :...location = Living Room(2): Undefined (20)
        location = Dining Room(25): Undefined (53.3)
        location = Kitchen(10):
        :...Small cupboard drawer top = on: Undefined (18.3)
            Small cupboard drawer top = off:
            :...Ceiling cupboard door middle = on: Undefined (10.2)
                Ceiling cupboard door middle = off:
                :...Floor cupboard drawer right = on: Undefined (10.2)
                    Floor cupboard drawer right = off:
                    :...Fridge = on: Undefined (81.5/26)
                        Fridge = off: Preparing_Hot_Drink (555.2/206.5)

Etc.
```

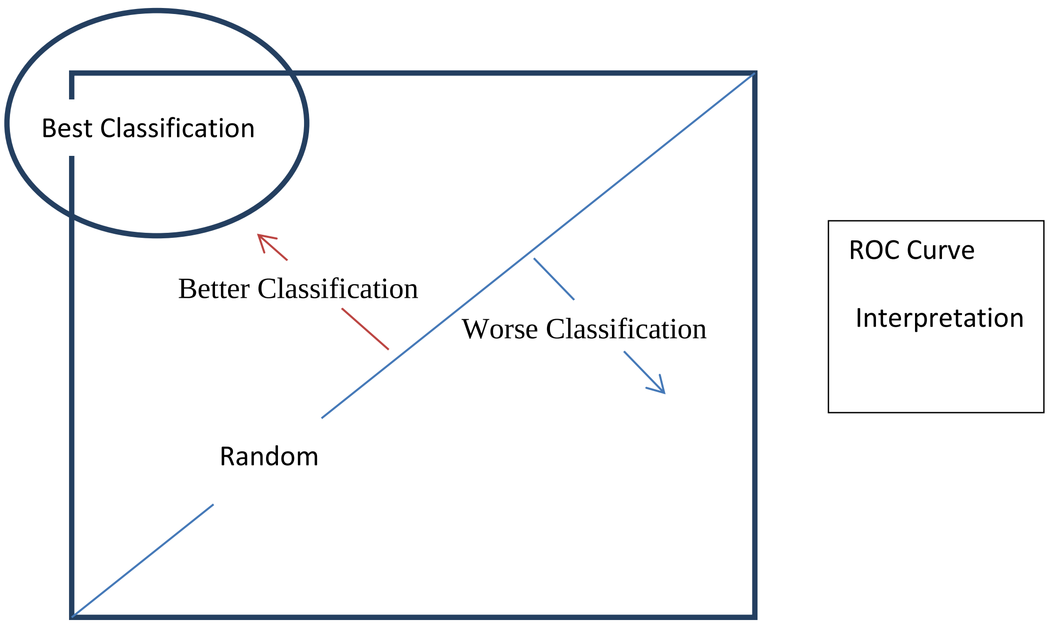
Figure 12: An example of the rule derived from the 'Preparing a hot drink' data.

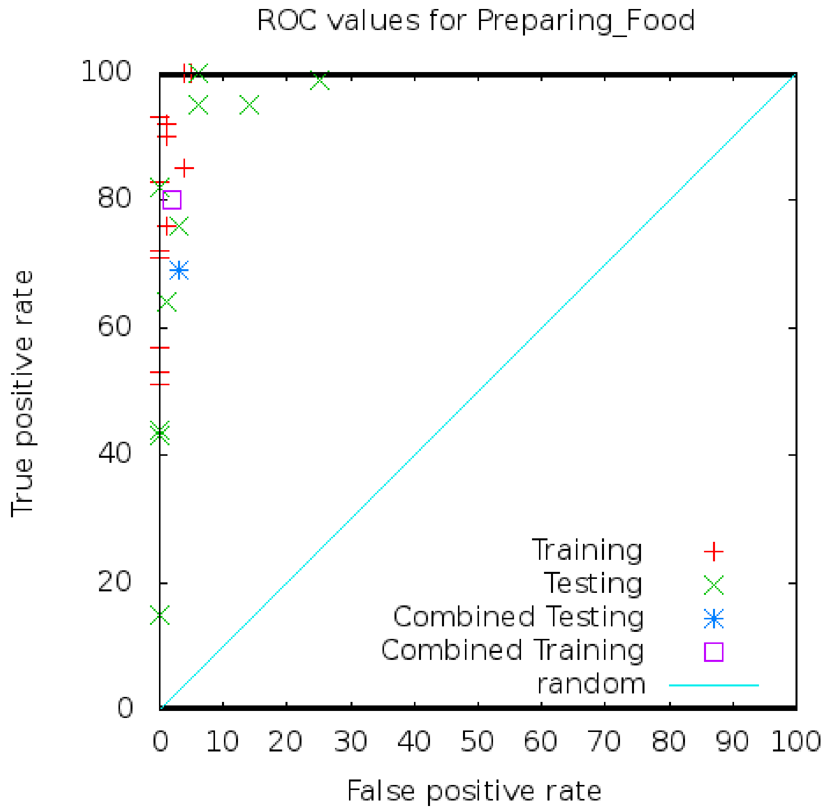
```
Evaluation on training data (1167 cases):

Trial       Decision Tree
-----     ----------------
         Size      Errors

   0        4    4( 0.3%)
   1        4  101( 8.7%)
   2        2   48( 4.1%)
   3        9   66( 5.7%)
   4        3   31( 2.7%)
   5        7   72( 6.2%)
   6        5    0( 0.0%)
   7       10  106( 9.1%)
   8        5    0( 0.0%)
boost              0( 0.0%)   <<


   (a)  (b)  (c)  (d)  (e)  (f)  (g)  (h)  (i)  (j)  (k)  (l)  (m)  (n)  (o)    <-classified as
  ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ---- ----
                                                                           (a): class Sitting_Living_Room
                                                                           (b): class TV_ON
                                                                           (c): class Watching_TV
                                                                           (d): class Preparing_Food
                                                                           (e): class Using_Toaster
                                                                           (f): class Preparing_Cold_Drink
                                189                                        (g): class Preparing_Hot_Drink
                                                                           (h): class Laying_Table
                                                                           (i): class Having_Meal_Living_Room
                                                                           (j): class Cleaning_Table
                                                                           (k): class Spare_Time_Living_Room
                                                                           (l): class Sitting_Dining_Area
                                                                           (m): class Computer_On
                                                                           (n): class Using_Computer
                                                                     978   (o): class Undefined
```

Figure 13: Boosting and Confusion Matrix for Preparing Hot Drink. In this example the application of the boosting algorithm ensures that 100% of the cases are identified correctly.

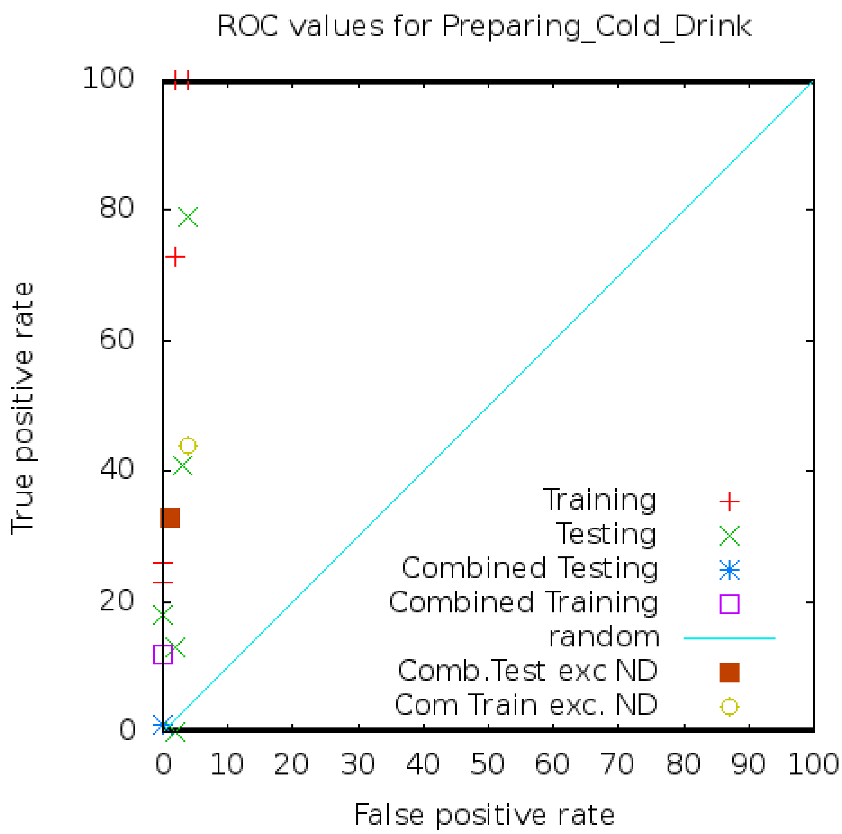
The ROC curves shown below illustrate the effectiveness of the learning algorithm in four cases, Preparing Food, Preparing Hot Drink, Preparing Cold Drink and Using Toaster.
Data points near the top left corner indicate a more effective classification. Cases to the left of the central 45 degree line (random) show a positive result – whereas those to the right of the line show a negative result.

True positives and false positive rates are shown on each axis. A high false positive rate would indicate that the system was consistently classifying an event *as* the task when it is in fact something else, a low true positive rate would indicate that the system was consistently mis-classifying it as something else when it was in fact the task.



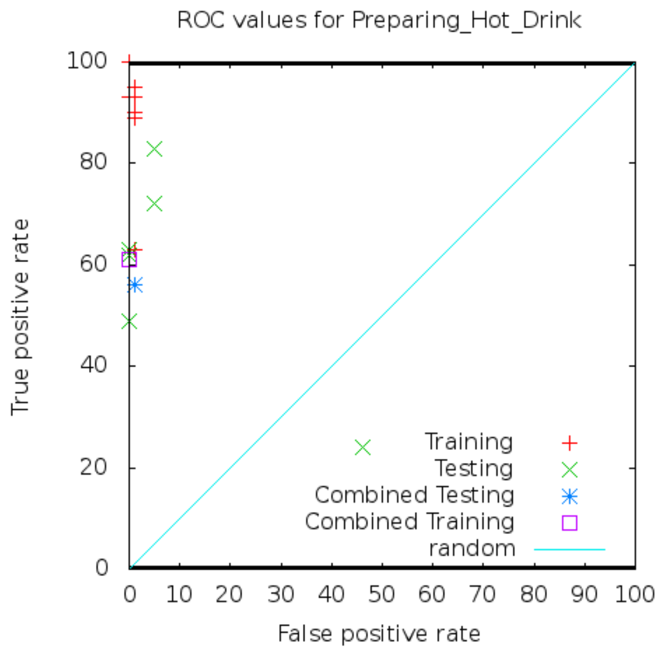


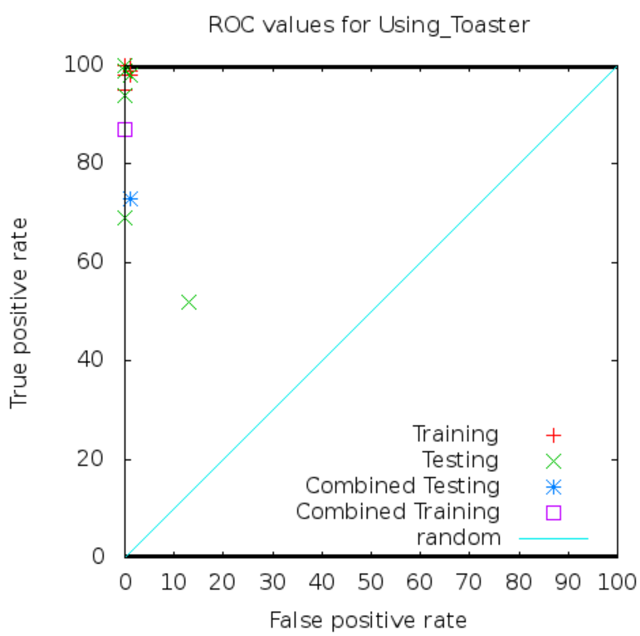
For the ‘Preparing Food’ example above each data point indicates a participant’s training or testing session. Also shown are the combined results after aggregation of data of all of the 14 participants into one data set. Overall clusters occur in the top left quadrant indicating a strong level of learning and recognition performance.



For the ‘Preparing Cold Drink’ example relatively low levels of learning and recognition performance were encountered. This, we believe, is due to two factors; firstly, that in some cases there was no training data and so the participant did not execute this task in training, but did subsequently carry it out in the open session. Secondly, the number of training instances overall for this task is low – people generally favoured making hot drinks rather than cold drinks. Two additional combined analyses are shown in the graph, both excluding instances where data was not present (the points labelled ND in the diagram above are where we excluded these cases). These show improvements over the raw data, however, the recognition rate is still relatively low.

For the ‘Preparing Hot Drink’ example a strong performance on the training data was not reflected in the test data with some outliers reducing performance. This was due to cases where there was no training data and so the participant did not execute this task in training, but did subsequently carry it out in the open session and thus overall reducing the performance.





As might be expected the 'using Toaster' example gave a strong performance on both the training and test data. This classifier is based on a single sensory event (i.e. the toaster being on) and therefore should give good classification results.

6.1.7 The end user as the 'observer/annotator'

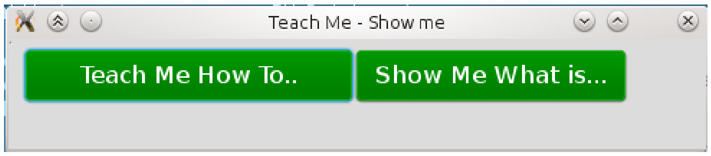
The results of the analysis above indicated that such a learning approach might allow the robot to recognize human activities in the robot house. However, in a 'real' situation we are faced with having no observer of human actions and no annotator of those actions to derive a classification set.

In order to provide a solution to this issue the obvious choice is to allow the house occupant to become the observer/annotator by simply informing the robot when tasks are starting and finishing.

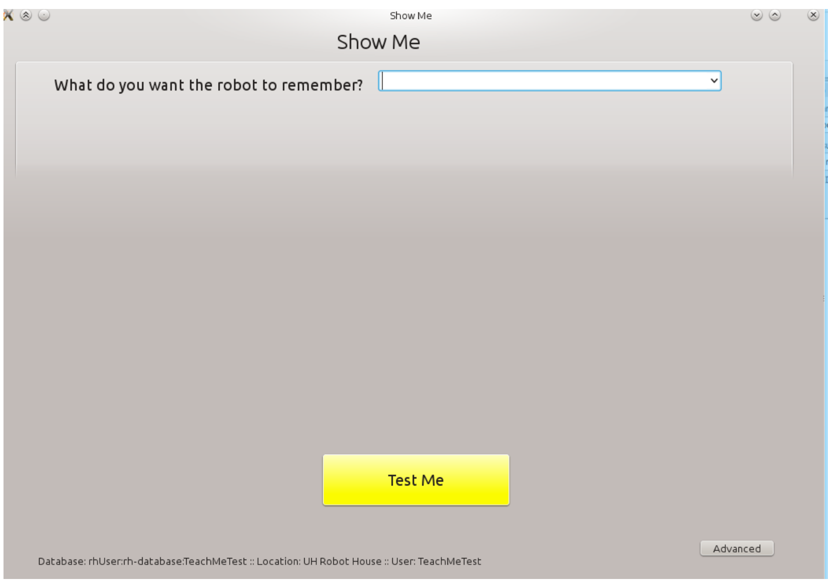
This is 'Show Me' part of the 'Teach Me/Show Me' interface.

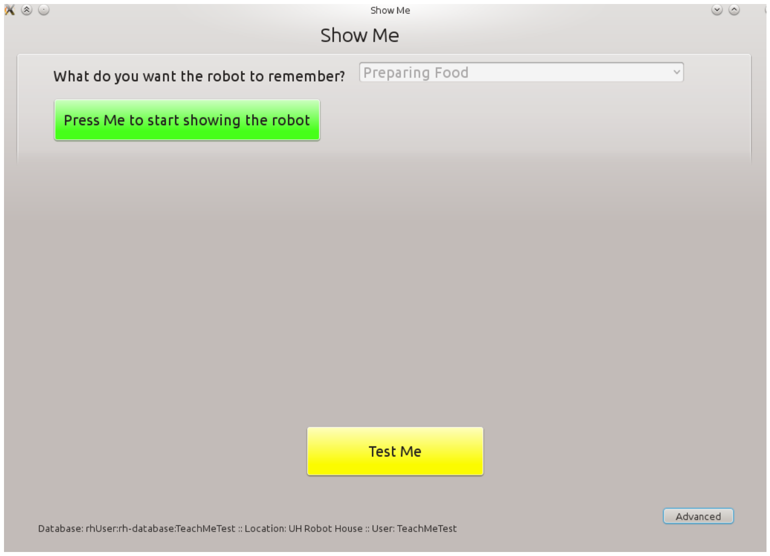
6.1.8 The 'Show Me' interface

The 'show me' interface consists of a simple GUI which appears when the 'Show Me What Is' button is pressed on the 'Teach Me/Show Me' interface.

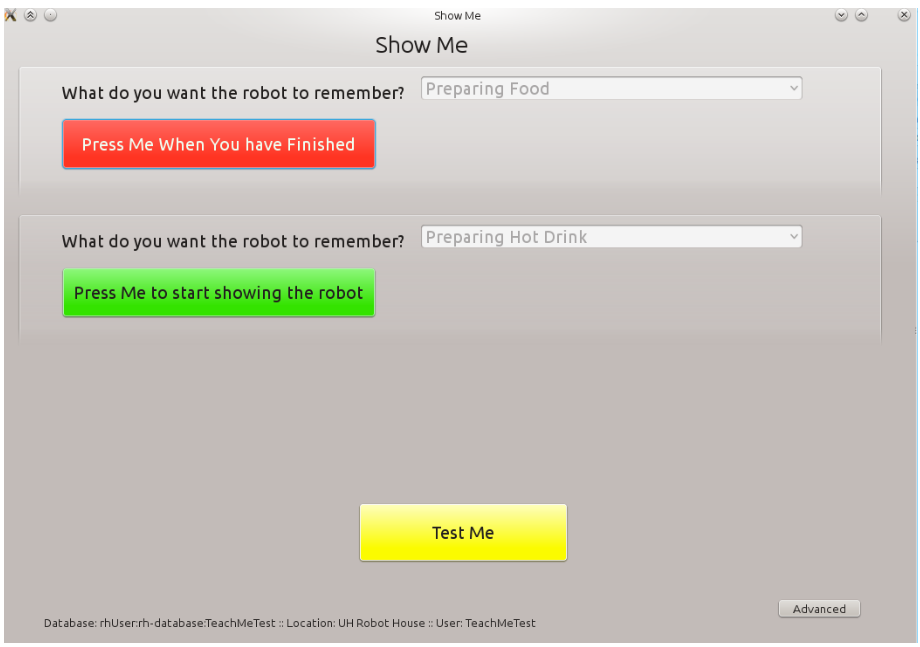


The user is presented with a screen which asks them to label their actions, and to press a start button when actually carrying out that action.



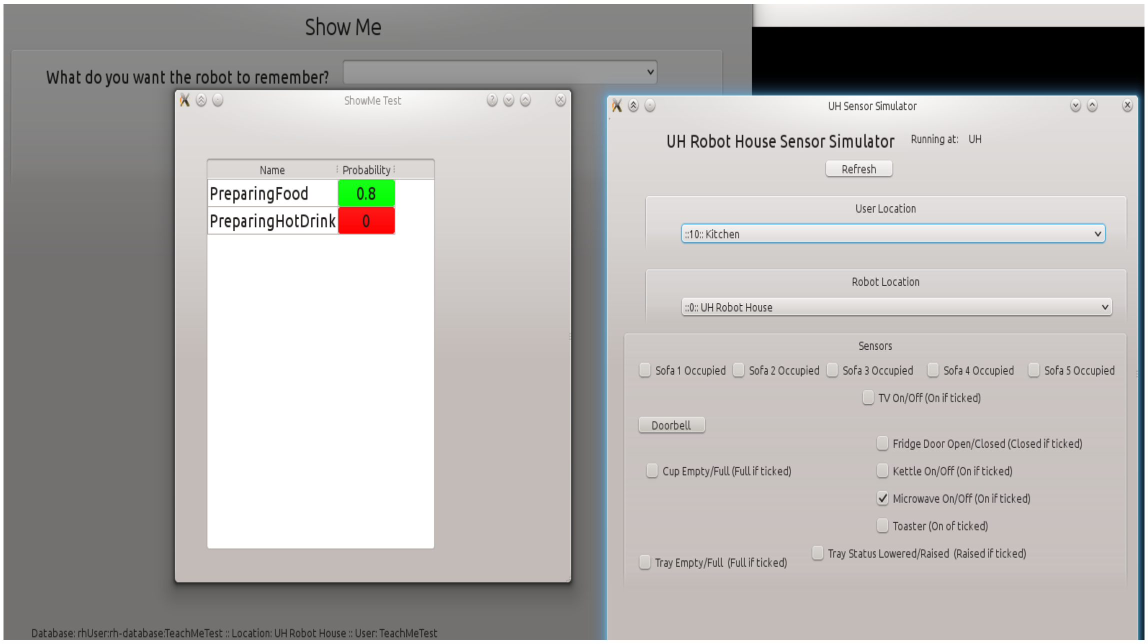


Here the user has entered 'Preparing Food' and when ready presses the 'press me to start showing the robot' button. They then carry out actions associated with preparing food (For example starting the microwave oven).



If a sub-task is required (in this case "Preparing a Hot Drink"), the user can continue to enter new tasks up to a maximum of 3 levels deep. Once each task completes, the user presses the red "Press me when you have finished" button.

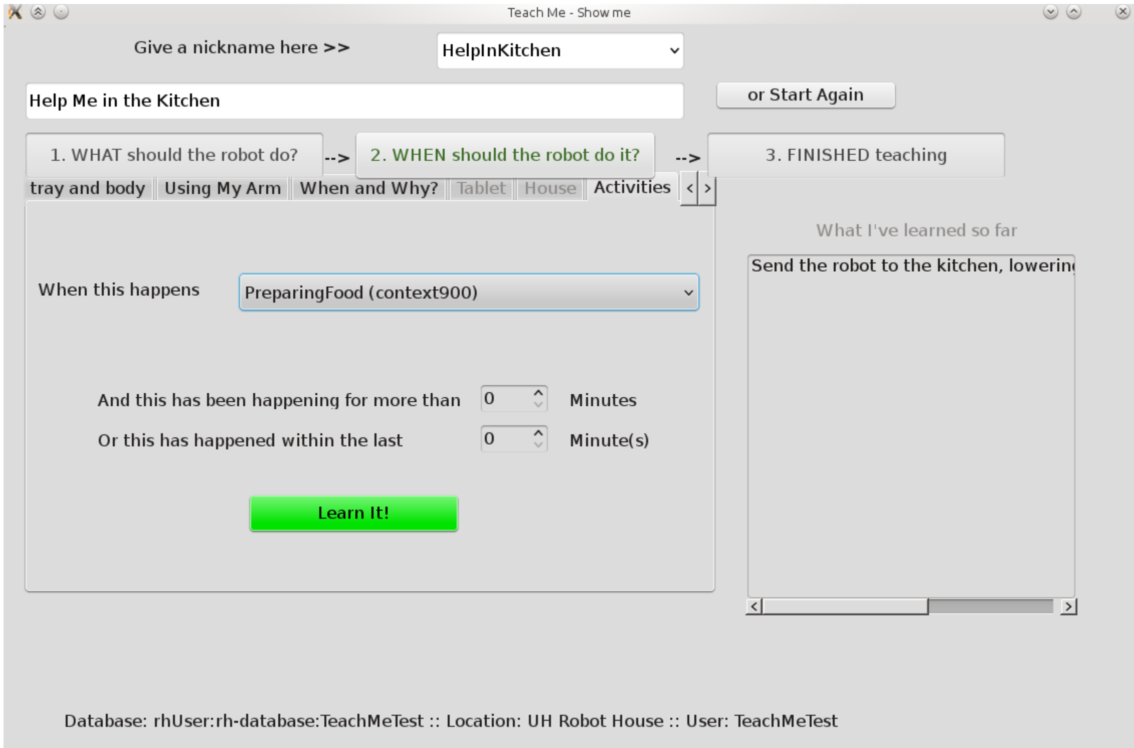
Testing can be carried out by pressing the 'Test Me' button. This operates in real-time and allows the user, whilst repeating the task, to check if the system correctly identifies it. A probability % is also given based on the predicted accuracy of the real-time classifier using the learned rules. The colour of the classifier symbol turns green if the probability exceeds 50%. Note that the system automatically creates lexical symbols which are then available within the robot teaching interface.



In the testing example (shown being tested with the house simulator), the microwave is on, therefore the system infers that ‘Preparing Food’ is 80% certain. However as the kettle is off ‘Preparing Hot Drink’ is very unlikely (0%).

The user now has the opportunity to use the ‘TeachMe’ system to teach the robot a new behavior based on what it has learnt.

For example, having taught the robot about ‘Preparing Food’, the user could create a behavior called ‘Help me in the kitchen’ which made the robot come to the kitchen whenever ‘Preparing Food’ was active. See the GUI screen below.



7 Initial Formative Evaluation of the 'Show Me' interface

A small evaluation of the learning system was carried out in addition to the long-term summative studies. The limited evaluation involved 3 persons, all of which were "informal" carers. They typically looked after an elderly relative.

All of the informal carers had previously been exposed to the 'Teach Me' system as part of the summative study and were therefore familiar with the teaching process.

7.1 Procedure

The experimenter explained the purpose of the study and ensured that they understood the instructions (see Appendix for participant instructions).

The experimenter then chose one of the activities in Table 10 and explained to the participant how to use the 'show me' GUI to allow the robot to learn about this activity – typically by actually carrying out that activity whilst the 'start showing me' button was active.

The participant were then asked to choose one of the other activities shown in Table 10 and use the 'show me' GUI to allow the robot to learn about the activity. They then tested this activity to ensure that the robot correctly identified it.

Show Me - Activities
Create an activity called ‘Watching TV’.
Create an activity called ‘Relaxing on the Sofa’.
Create an activity called ‘Preparing a hot drink’.
Create an activity called ‘Preparing a Ready Meal’ using the microwave.
Create an activity called ‘Kitchen Activities’ which is active when ‘preparing a hot drink’ or ‘Preparing a ready meal’ or when any other kitchen activity is being carried out.

Table 10: The participant chooses an activity to show to the robot. For example, if ‘Preparing a ready meal’ was chosen the participant would enter this ‘Kitchen Activities’ into the ‘ShowMe’ interface and press the start button. They would then typically open the fridge, pick up the read meal and place it in the microwave. At this point they would enter the text ‘Preparing a ready meal’ into the show me interface and press the start button on the GUI. They would then turn on the microwave, and when complete remove the meal. At this point the ‘Show Me’ stop button would be pressed on both the ‘Kitchen Activities’ and the ‘Preparing a ready meal’ GUI entries.

Having successfully tested that the robot had learned about this activity, the participant was asked to choose the corresponding teaching task in Table 11. For example, if the robot had learned about ‘watching TV’, then the behavior involving ‘watching TV’ would be chosen.

The participant then taught the robot the chosen behavior and subsequently tested that it worked. For example, for ‘watching TV’, that the robot would approach the participant (who was now sitting on the sofa watching TV) and inform them about an upcoming TV program.

Following this the participant was asked to complete the System Usability questionnaire.

Teach Me – Tasks
Teach the robot that if it is 7.30 and you are 'watching TV' then remind you that your favourite program is about to start.
If you have been 'Relaxing on the sofa' for more than 30mins make the robot come to the sofa and tell you to 'move about a bit'
If there are 'Kitchen Activities' make the robot come to the kitchen and offer to carry the drink or meal to the dining table.

Table 11: Having shown the robot a typical set of 'Kitchen Activities', this symbol becomes available in the robot teaching system and the robot can choose this higher level semantic 'sensor' directly.

7.2 Results from Limited Formative Evaluation of the 'Show Me' interface

Note that the number of participants in the Show me trial was very low and at this stage results should be treated with caution.

7.2.1 System Usability Scores

The SUS scores for the *Show me* interface ranged from 67.5 to 80. The mean score was 75.83 and the median score was 80. This was larger than the expected average of 68 (Not significant: ($t(8)= 1.88$ $p= 0.2$)).

7.2.2 Ad-hoc Likert Scales

The following Likert Items were also given to the participants:

Question	Scale
Do you think it is useful teach a robot about your activities?	Useful - Not useful
Do you think that *you* would be able to teach the robot?	Yes - No
Would you be willing to teach the robot for someone else e.g. if you were a relative or carer of the other person?	Yes - No

Do you think that activities should already have been completely setup by someone else? Yes - No

The responses to these questions can be found below:

Question			Frequencies		
	Very Useful	Useful	Neither useful nor useless	Not Useful	Not Useful at all
Useful to teach robot activities	3	0	0	0	0
	Definitely Yes	Yes	Don't know	No	Definitely No
Able to teach robot activities	2	1	0	0	0
	Definitely Yes	Yes	Don't know	No	Definitely No
Teach robot for others	2	1	0	0	0
	Definitely Yes	Yes	Don't know	No	Definitely No
Activities completely set up	0	0	0	3	0

7.2.3 Open-ended Questions

The open-ended responses are shown below:

Comments on the Show me interface
I think it is a great idea to personalise the robot for an individual's needs. But also think this can be used alongside prepared repetitive tasks. I think also very important for the robot to learn activities rather than or as well as one off tasks. When teaching activities need to show robot in simple exaggerated steps so that it does not confuse activities
Not completely set up but a range of everyday types of activities which can be personalised.

One of the participants argued that a similar approach be adopted as for the Teach me feature, in that the system comes with some activities pre-set but that these can be personalised. The other participants reiterated the need for such a feature.

7.3 Summary of Results

Clearly such a small sample may only be indicative, however the results from the SUS suggested that participants found the interface relatively easy to use.

The three participants all found the Show me feature useful, and felt confident in their ability to use a feature like this to teach a robot about their own activities or to use on behalf of someone else. They also felt that this should not be something that was already set up prior to use.

8 Discussion and Future Work

In this document we have presented formative evaluations of the robot teaching system ("teach me"), the memory visualisation system and an initial evaluation of the robot learning system ("show me").

The results of these studies indicate that such facilities would be readily accepted for use by carers, relatives and the elderly themselves. However, with increasing age, the willingness to learn new ways to operate, by personalising a robot's behaviours, decreases.

The exploration, via the 'show me' system, of creating higher level semantics, is we believe a novel and promising way to ease the teaching burden. For example, being able to instruct a robot by using everyday terms, such as "when it's time for bed do …" or "if I'm making dinner do …". We have only partially explored such opportunities and issues which surround 'showing' a robot typical activities and this work is at an early stage. A number of improvements and enhancements to such facilities would be to firstly, use both inductive and *predictive* mechanisms to increase the reliability of the robot recognising user activities.

Prediction algorithms already exist which use past sensory data to predict possible next actions (e.g. (Gopalratnam & Cook 2003)) and some preliminary work has been carried out in this respect (see the Appendix for details of this work). An extension of this work would be to use those predictions to then predict again – effectively creating a predictive forward model for the robot. This forward model then being subject to the inductive algorithm, which would now use both historical and predicted sensor vectors to make a decision of user activity. Secondly, to improve the interfaces for both the "teach me" and "show me" functionalities by interface design experts. Currently these interfaces have been designed by non-interface experts, and can appear rather unintuitive.

The opportunities for exploring episodic episodes related to user activities have been explored via the visualisation system, however much more research in this area is required. For example, although users can mark particular episodes as important, the system itself cannot automatically carry this out. Further work relating past episodes and their importance to future episodes may be one direction where fruitful research could operate.

Finally, we have demonstrated in this work that personalisation of an autonomous robot is possible in a domestic environment. Further analysis of the exploitation and commercialisation of these findings may also be a positive next step.

References

Brooke, J. (1996), ‘SUS-A quick and dirty usability scale’, *Usability evaluation in industry* **189**, 194.

Chen, C., Das, B. & Cook, D. (2010), A data mining framework for activity recognition in smart environments, *in* ‘Proc. International Conference on Intelligent Environments’.

Das, S., Cook, D. J., Bhattacharya, A., Heierman, E. & Lin, T. (2002), ‘The role of prediction algorithms in the MavHome smart home architecture’, *IEEE Personal Wireless Communications* **9**(6), 77–84.

Duque, I., Dautenhahn, K., Koay, K. L., Willcock, I. & Christianson, B. (2013), Knowledge-driven user activity recognition for a smart house. development and validation of a generic and low-cost, resource-efficient system, *in* ‘In ACHI 2013, The Sixth International Conference on Advances in Computer-Human Interactions’, pp. 141–146.

Feder, M., Merhav, N. & Gutman, M. (1992), ‘Universal prediction of individual sequences’, *Transactions on Information theory* **38**(4).

Federosky, E., feder, M. & Weiss, S. (1998), Branch prediction based on universal data compression algorithms, *in* ‘Proceedings of Annual Symposium of Computer Architectures’.

Gopalratnam, K. & Cook, D. J. (2003), Active lezi: An incremental parsing algorithm for sequential prediction, *in* ‘In Sixteenth International Florida Artificial Intelligence Research Society Conference’, pp. 38–42.

Gosling, S. D., Rentfrow, P. J. & Swann Jr, W. B. (2003), ‘A very brief measure of the big-five personality domains’, *Journal of Research in personality* **37**(6), 504–528.

Gu, T., Wu, Z., Tao, X., Pung, H. & Lu, J. (2009), epsICAR: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition, *in* ‘Proc. of the International Conference on Pervasive Computing and Communications’, pp. 1–9.

Koay, K. L., Syrdal, D. S., Walters, M. L., Oskoei, M. A. & Dautenhahn, K. (2014), ‘Social roles and baseline proxemic preferences for a domestic service robot’, *International Journal of Social Robotics* .

Krishnan, N. C. & Cook, D. J. (2014), ‘Activity recognition on streaming sensor data’, *Pervasive and Mobile Computing* **10, Part B**(0), 138 – 154.

Lester, J., Choudhury, T., Kern, N., Borriello, G. & Hannaford, B. (2005), A hybrid discrimative/generative approach for modelling human activities, *in* ‘Proc. of the 19th International Joint Conference on Artificial Intelligence (IJCAI ’05)’, pp. 766–772.

Quinlan, J. R. (1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Mateo, CA.

Rashidi, P. & Mihailidis, A. (2013), ‘A survey on ambient-assisted living tools for older adults’, *Biomedical and Health Infomatics* **17**(3), 579–590.

Reiser, U., Connette, C., Fischer, J., Kubacki, J., Bubeck, A., Weisshardt, F., Jacobs, T., Parlitz, C., Hagele, M. & Verl, A. (2009), Care-o-bot creating a product vision for service robot applications by integrating design and technology, *in* ‘Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on’, IEEE, pp. 1992–1998.

Saunders, J., Burke, N., Koay, K. L. & Dautenhahn, K. (2013), A user friendly robot architecture for re-ablement and co-learning in a sensorised homes, *in* ‘Proc. 12th European Conf. Advancement Assistive Technology in Europe, (AAATE13)’.

Saunders, J., Nehaniv, C. L. & Dautenhahn, K. (2006), Teaching robots by moulding behavior and scaffolding the environment, *in* ‘Proceedings of the ACM Conference on Human-Robot Interaction (HRI06)’, ACM Press, pp. 118–125.

Vitter, J. & Krishnan, P. (1996), ‘Optimal prefetching via data compression’, *Journal of the ACM* **43**(5), 771–793.

Ziv, J. & Lempel, A. (1978), ‘Compression of individual sequences via variable rate encoding’, *Transactions on Information Theory* **24**, 530–536.

Appendix

Participant Instructions for the Teaching Formative Study (Teach Me)

1) We would like you to teach our robot some new tasks.

   You can do this using a special program called TeachMe.

   The experimenter will first show you how to use this program to create a robot task.

   We will then test whether the robot has understood the instructions and whether it carries out the task successfully.

2) It will then be your turn to teach the robot some new tasks.

   You can choose a task from the table below (see Robot Tasks 1).

   When you have taught the robot a task you will then test it together with the experimenter.

3) If you have time you can teach the robot another task.

   This time teach the robot a task from the table labelled ‘Task type 2’.

   And when complete, test again.

   If you still have time, then repeat choosing a task from the table ‘Task type 3’ and when complete test again.

4) After this we will ask you to fill in a questionnaire and then one of the experimenters will ask you a few questions about your experience teaching

the robot.

Please remember that we are testing the robot and the teaching system, we are NOT testing you!

Thank you for your time and help!

Participant Instructions for the Learning Formative Study (Show Me)

1. We would like you to show our robot the sort of things you usually do around the house.

    This is so that the robot can learn your daily routines and activities.
    You can then use the TeachMe program to personalize the robot's behavior.
    To allow the robot to learn activities we use the 'ShowMe' program.
    The experimenter will first show you how to use this program to make the robot learn a new activity.
    We will then test whether the robot has understood and whether it recognizes the activity when you repeat it.

2. It will then be your turn to show the robot some new activities.

    You can choose an activity from the table below (see Show Me -Activities).

    When you have taught the robot an activity you will then test it together with the experimenter. You can then teach the robot by referring to this new activity.

    Teach the robot a task from the table labelled 'Teach Me - Tasks.
    And when complete, test again.

3. After this we will ask you to fill in a questionnaire and then one of the experimenters will ask you a few questions about your experiences of showing the robot new activities.

Preliminary Work on Sequential Activity Prediction

During the final part of the ACCOMPANY project investigations into the use of sequential prediction algorithms to predict possible user activities was carried out. This initial study was based on the data set used by Duque et al. (2013). Sequential activity prediction attempts to use historical data (typically held as a time-series of sensory states) to predict the next sensory state given the current state. The idea of building a predictor from deterministic sequences was explored by Feder and colleagues (Feder et al. 1992) who proved that the Markov predictors based on the well-known LZ78 algorithm (Ziv & Lempel 1978), versions of which are known as 'zip' compression, attained optimal predictability. This attribute has been used for various applications including computer program branch prediction (Federosky et al. 1998) and memory page pre-fetching (Vitter & Krishnan 1996).

An extension of the algorithm called 'ActiveLezi' was created by (Gopalratnam & Cook 2003) and subsequently used as a predictor in smart home environments. In the work presented here we have implemented the ActiveLezi algorithm on the data sets derived from the UH Robot House by Duque to assess its feasibility and possible future usage.

ActiveLezi is based on LZ78 but addresses some shortcomings in relation to information phase crossing boundaries by using a sliding window to collect statistics on all possible contexts. For a detailed explanation of ActiveLezi see (Gopalratnam & Cook 2003) and for and explanation of LZ78 see (Ziv & Lempel 1978).

ActiveLezi was used as a predictor of actions in the MavHome smart home project at the University of Texas at Arlington (Das et al. 2002) and the results indicated the usefulness of such an approach.

Applying ActiveLezi to Data from the UH Robot House

The ActiveLezi algorithm was implemented and the input procedures modified to use data collected from the UH Robot House. This is exactly the same data as shown in section 6.1.4 and based on 14 participants carrying out two sets of equivalent activities in the robot house. For this study the training set was used to build the ActiveLezi predictor model and the testing sets were used to test the model. Note that the data labelling is not necessary in this context as we are attempting, at this early stage, to simply predict the next sensory event.

Results

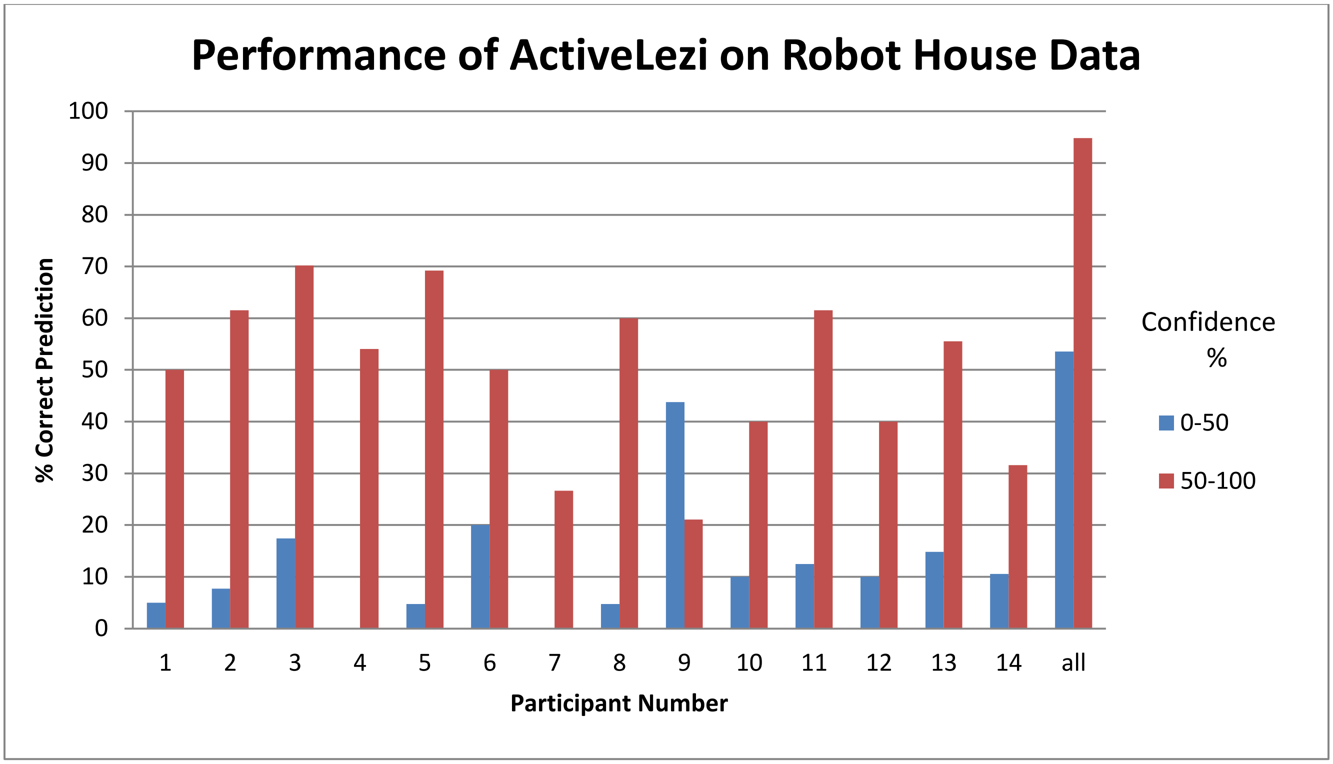


Figure 14: Each participant is shown with two measures. The blue bar indicates the percentage correct predictions when ActiveLezi confidence was less than 50%. The red bar indicates the percentage correct predictions when ActiveLezi confidence was greater than 50%. The final columns (marked 'all') shows the result of combining all of the participants data.

ActiveLezi provides a confidence level on each of its predictions. The graph in Figure 14 shows, for each participant, two measures of confidence. The first, shown in blue, is where ActiveLezi predicts the results with a confidence of less than 50%. The red bars show a confidence level greater or equal to 50%.

Where ActiveLezi predicted with a confidence of greater than 50%, in the majority of cases the actual prediction was correct in over 50% of the time. Correct prediction levels increased significantly when all of the participant data was combined reaching a level of around 95% correct predictions.

A more in depth investigation of the sensory events predicted by ActiveLezi indicated that the algorithm was making good predictions, however the usefulness of the actual prediction was questionable. For example if the current sensor status indicated that the cupboard was open, the probable next sensory occurrence may be that the cupboard would be closed. This is exemplified in Figure 15 .

Therefore, further work on this area of prediction, possibly to use the current prediction to predict the next state, and subsequently use these forward predictions to verify the rule induction algorithm, may be a useful next step.

```
 Actual event: Toaster  OFF *** Correct!!
Predicted: Fridge 0.94638 ON
          Actual event:  Fridge  ON *** Correct!!
Predicted: Fridge 0.999611 OFF
          Actual event:  Fridge  OFF *** Correct!!
Predicted: Chair seatplace 0 0.948043 ON
          Actual event:  Chair seatplace 0  ON *** Correct!!
Predicted: Chair seatplace 0 0.985678 OFF
          Actual event:  Chair seatplace 0  OFF *** Correct!!
Predicted: Chair seatplace 0 0.987532 ON
          Actual event:  Chair seatplace 0  ON *** Correct!!
Predicted: Computer Dining Area 0.971288 ON
          Actual event:  Computer Dining Area  ON *** Correct!!
Predicted: Chair seatplace 0 0.969748 OFF
          Actual event:  Chair seatplace 0  OFF *** Correct!!
Predicted: Cupboard small drawer top 0.987068 ON
          Actual event:  Cupboard small drawer top  ON *** Correct!!
Predicted: Cupboard small drawer top 0.997665 OFF
          Actual event:  Cupboard small drawer top  OFF *** Correct!!
```

Figure 15: The output from the ActiveLezi sequential Predictor. The predicted line indicates what ActiveLezi is predicting after the last sensory even and the confidence level. The actual event is the real next event. If both the predicted and actual events match the line is marked with 'Correct'.

A sample of the output from ActiveLezi is shown in Figure 15. Starting from the top of the figure, the last real event was the toaster being turned off. ActiveLezi then predicts that the fridge will be opened with 95% confidence. The actual next event is indeed the fridge being opened. ActiveLezi then predicts the fridge will be closed with a confidence of almost 100%. The next real event is the fridge being closed. This sample is taken from the combined participant data.

System Overview Diagrams For TeachMe and ShowMe systems

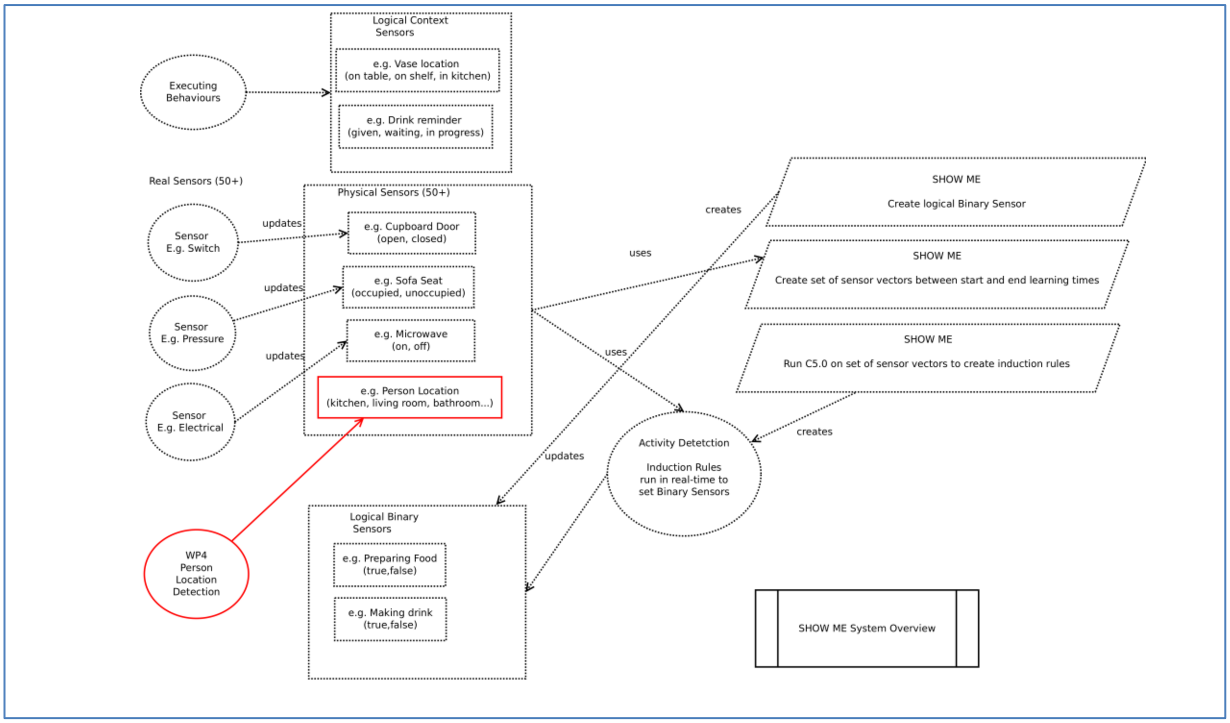


Figure 16: Show Me System Overview

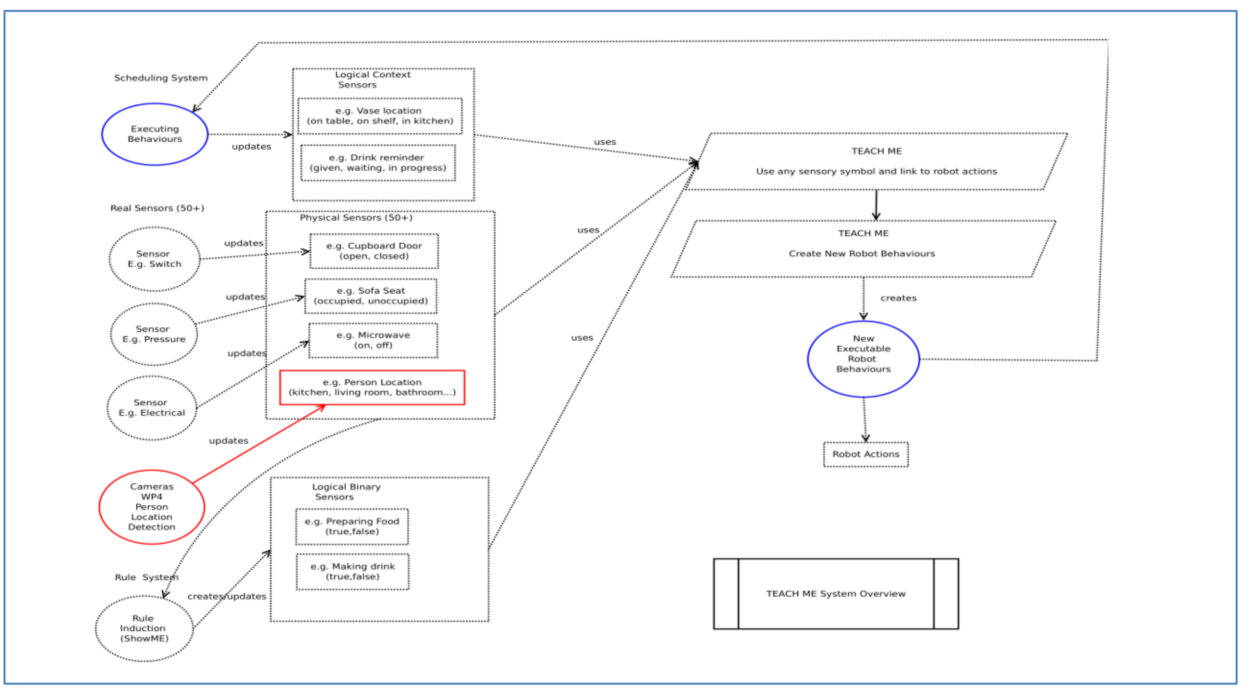


Figure 17: Teach Me System Overview