



FP6-IST-002020

COGNIRON

The Cognitive Robot Companion

Integrated Project

Information Society Technologies Priority

D6.2005

**Joint RA6 Deliverable
Intentionality and Initiative**

Architecture, Decision and Human-Robot Interaction

Due date of deliverable: 31/12/2005

Actual submission date: 25/01/2006

Start date of project: January 1st, 2004

Duration: 48 months

Contributing partners:

LAAS, IPA, UniBi, UH, UKA

Revision: final

Dissemination Level: PU

Executive Summary

This document reports on activities, advances and results related to RA6.

We have conducted investigations for the design of a conceptual architecture for a robot companion that acts, learns and interacts with humans. This architecture and its main features are described.

Concerning decisional issues involved in human-robot interaction (HRI), we have devised a generic framework that provides a basis for a principled way to deal with robot task achievement in presence of humans or in synergy with humans.

We have conducted user studies regarding subjects' perceptions of robots and the relationships subject/robot personalities. Their analysis has allowed to emphasize issues that will serve as guidelines in the synthesis of robot tasks and motions.

We have also made substantial progress in the implementation of architectural components: a robot supervisor derived from the HRI framework, a task representation that is suited for task adaptation and extension at run-time, a refinement of the BIRON robot architecture.

Besides advances in the robot architecture and HRI related decisional processes, we have produced software components and tools that are used in the three Key Experiments.

Role of RA6 in Cogniron

A cognitive robot companion must exhibit capabilities to understand its environment (RA5), to learn (RA4), to interact with people (RA1, RA2, RA3) and to make decisions. In RA6, a conceptual architecture is studied that provides a framework that integrates all these capabilities. The conceptual architecture is to be implemented partially in the Key Experiments (RA7). The decision-making capability is also more specifically studied in RA6, be it for autonomous deliberation and task achievement, or for human-robot collaborative problem solving. When interacting with a robot, people tend to attribute intentions to it according to its behavior and other factors. Studies on intentionality attribution are also conducted within this RA.

Relation to the Key Experiments

The issue of architecture is studied in RA6, but is also related in the project to the implementation of the Key Experiments (RA7) and to integration of other RAs. The architectural design is instantiated in the various key experiments. The collaborative problem solving framework as well as intentionality attribution is related to KE1 and KE2.

Contents

1	Introduction	5
2	A Conceptual Architecture for a Cognitive Robot	6
2.1	Introduction	6
2.2	Cognitive Architectures	6
2.2.1	SOAR	7
2.2.2	ACT-R	8
2.2.3	Conclusion on Cognitive Architectures	9
2.3	Robot Control Architectures	9
2.3.1	Robot Control	9
2.3.2	Examples of Robot Control Architectures	10
2.3.3	Conclusion on Robot Control Architectures	15
2.4	A Conceptual Architecture for Cogniron	15
2.4.1	Requirements and Architecture	15
2.4.2	Cognitive capacities	17
2.5	Future Work	19
3	A Framework for Human-Robot Collaborative Problem-Solving	19
3.1	Context	19
3.2	A Decisional framework	20
3.3	Human-Aware Task Planning	22
4	Models and Algorithms for cognitive robots supervision	23
4.1	Design considerations	23
4.2	Supervision system	28
4.3	Results	33
4.4	Future work	33
5	Advances in Robot Control Architecture Implementation	34
5.1	BIRON architecture	34
5.2	Control Architecture tools related to “Programming by Demonstration”	34
5.2.1	Developments	35
5.2.2	Future work	36
5.3	LAAS Open Software for Autonomous Systems	36
6	Intentionality Attribution	38
6.1	Overview	38
6.2	In depth evaluation of the relationship between subject and robot personality and subjects	38
6.3	The association between peoples’ personality and spatial-comfort zones during human-robot interaction trials	40
6.4	Human comfort levels during human-robot interaction trials	41
6.5	How should a robot attract a person’s attention? Preferences toward different robot appearances, gestures and cues (new user studies in 2005)	42
6.6	How can a robot attract people’s attention?	43
6.7	Future work	44

7 Future Work	44
8 References	45
9 References	45
9.1 Applicable documents	45
9.2 Reference documents	46
10 Annex: papers related to RA6	50

1 Introduction

Research activity 6 deals with the decision-making mechanisms that are induced by the demanding context of personal robots.

Three complementary research streams have been defined:

- A control architecture for a robot companion that acts, learns and interacts with humans.
- Schemes for interactive human-robot problem solving.
- Attribution of intentionality to the robot, and its use in human-robot decisional interaction.

The work in RA6 for the second workplan has essentially pursued in the same research in terms of basic concepts and approaches but with the perspective of the key experiments that have been used as an opportunity to confront to real problems and to integrate with the other RAs.

The work was conducted through four workpackages:

- WP6.1 A generic architecture for a cognitive robot
- WP6.2 Models and Algorithms for cognitive robots supervision
- WP6.3 Human-Robot Collaborative Problem-Solving
- WP6.4 Intentionality Attribution

In order to ease legibility of the work achieved so far, the deliverable is organized in a way that gathers issues that are tightly connected:

- Section 2 reports on the research that we have conducted towards a conceptual robot architecture (contribution of WP6.1)
- Section 3 presents the decisional framework for human-robot interactive task achievement that we have devised (contribution of WP6.2, WP6.3 and WP6.4)
- Section 4 reports on the design and implementation of a robot supervision system derived from the framework presented in the previous section (contribution of WP6.2 and WP6.3)
- Section 5 reports on the development of robot architecture components and tools in relation with dialog and with learning (contribution of WP6.1, WP6.2)
- Section 6 reports on studies regarding subjects' perception of robots and the relationships subject/robot personalities (contribution of WP6.4).

2 A Conceptual Architecture for a Cognitive Robot

2.1 Introduction

A robot system architecture organizes the *components* of the system which are - at several levels of granularity - the functions the robot will be able to achieve. It describes how these components are linked and interleaved together, how they interact, what information they exchange, and at which temporal rate. In addition to the particular algorithms implemented in the components, the architectural design will hence determine the robot's capacities to achieve tasks and to react to events.

The architecture can be described at different levels, and from two different viewpoints:

- The conceptual architecture that describes the system in terms of functionalities and their interactions.
- The software architecture that will implement the concept, and by which it will be possible to actually achieve the integration of the developed algorithms (often implemented by different people).

A conceptual architecture may have different software implementations, according to the state of technology. In this document we will address both aspects.

From the general backgrounds of robotics and of cognitive science, we can identify four main capabilities that the architecture of a cognitive robot should enable, in addition to the more "basic" functions such as sensing, locomotion or simple manipulation, commonly present in a robot. Those are:

Spatio-temporal cognition: object, scene and situation modeling, recognition and interpretation

Decision-making, anticipation and reflexivity: ability to focus attention, to make choices, to plan in advance a course of action, to anticipate the results of these actions or of observed events, to reflect on own abilities and future.

Learning: improving abilities with experience, acquiring new abilities, building new knowledge from acquired information. Note that spatio-temporal cognition is also intricately related to learning.

Interaction and communication: to understand multi-modal commands (taskability of the robot), express intentions, explain actions.

In this document we will first overview some concepts and designs from the cognitive architectures literature (section 2.2) and from the robotics literature (section 2.3). Then we will discuss the requirements for the architecture of a cognitive robot companion and propose an architectural concept for Cogniron (section 2.4). In this document we keep a conceptual view. This concept is not going to be implemented as such in a unique architecture within the project. We shall rather develop partial instances in the key experiments. The architecture has indeed started to be mapped into actual implementations in the key experiments; this is presented on the RA7 deliverable.

2.2 Cognitive Architectures

The term "cognitive architectures" is commonly used in the AI and cognitive sciences communities to designate the organization of systems designed to model the human mind. Two of them, SOAR and ACT-R, are large long term projects and claim objectives of generality.

2.2.1 SOAR

SOAR (State, Operator And Result) [46, 34] is a system aiming at modelling human cognition. It is a long term project developed mainly at the University of Michigan and aiming at implementing instances of an artificial cognitive system in several domains, ranging from robotics to air combat. It is based on Newell seminal work on theories of cognition [40]. The authors of SOAR insist on the importance of system architecture.

Operational knowledge in SOAR is represented by production rules. To achieve a goal, the rules' conditions are matched to a "working memory" which contents is coded as sets of {attribute-values}. The working memory includes different kinds of knowledge. Current goals, states and operators are in a context stack in this memory. Hence SOAR implements the concept of "problem spaces" introduced by Newell. Rule consequences are actions that modify a preference memory. The operation of SOAR to achieve a goal will produce subgoals and therefore a goal hierarchy. Soar performs no conflict resolution between competing productions- all productions which match the current working memory fire.

Learning in SOAR is mainly based on a mechanism called "chunking" (other mechanisms such as reinforcement learning are being added). This process is similar to identifying macro-operators, i.e. new rules that abstract the succession of rules selected to achieve a goal. Therefore this new rule will be a new operational knowledge that is more adapted to the goal achievement, avoiding the impasse situations that were previously encountered.

Figure 1 depicts SOAR's architecture.

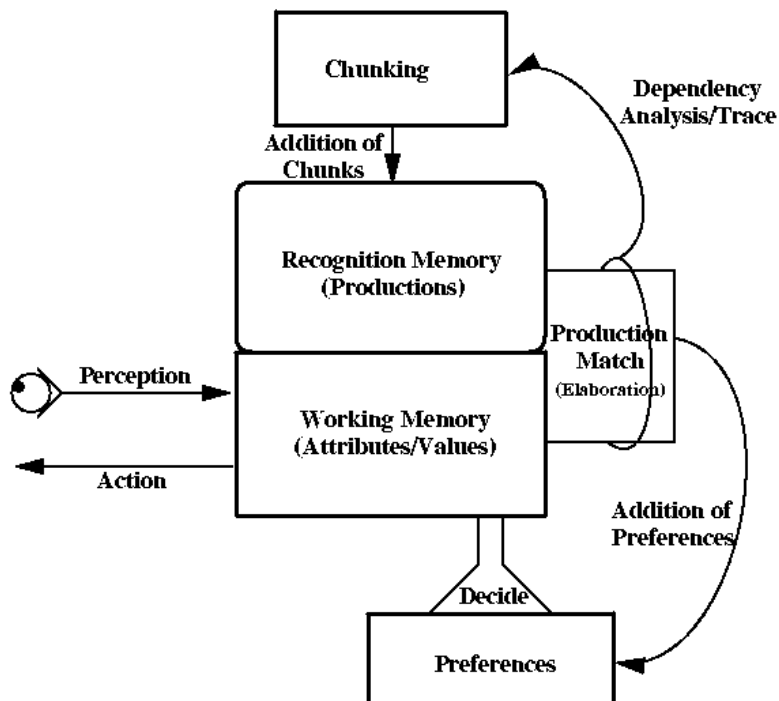


Figure 1: The high-level architecture of SOAR (from the SOAR website <http://sitemaker.umich.edu/soar>).

2.2.2 ACT-R

ACT-R (Adaptive Control of Thought-Rational) is a cognitive architecture proposed by John Anderson and colleagues, in development over several years, that represent for its authors a model of the human mind [16].

The general concept in ACT-R is a classical rule based system. Knowledge is organized in a declarative memory about facts and events and their relationships (the data structures in this memory are called "chunks" but the meaning is quite different from the chunks used in SOAR), and a set of production rules and procedures. The rules are selected according to their matching the chunks in the memory, and to their cost and probability of success. Their actions change the declarative memory or trigger actions in the world.

However, there is in addition a learning process. The declarative chunks have indeed an associated "base level" which increases according to their past selection. This results in using chunks that have already been selected, i.e, that were used in more successful activations of the rules. The costs and success rates of the rules is modified according to the outcome of their execution. This provides for an improvement of the global behavior. Furthermore, there is a "compilation" process that produces new rules from analyzing the chunks involved in goal achievement.

Figure 2 depicts the information flow in ACTR and the way the authors map the architecture to the brain structure.

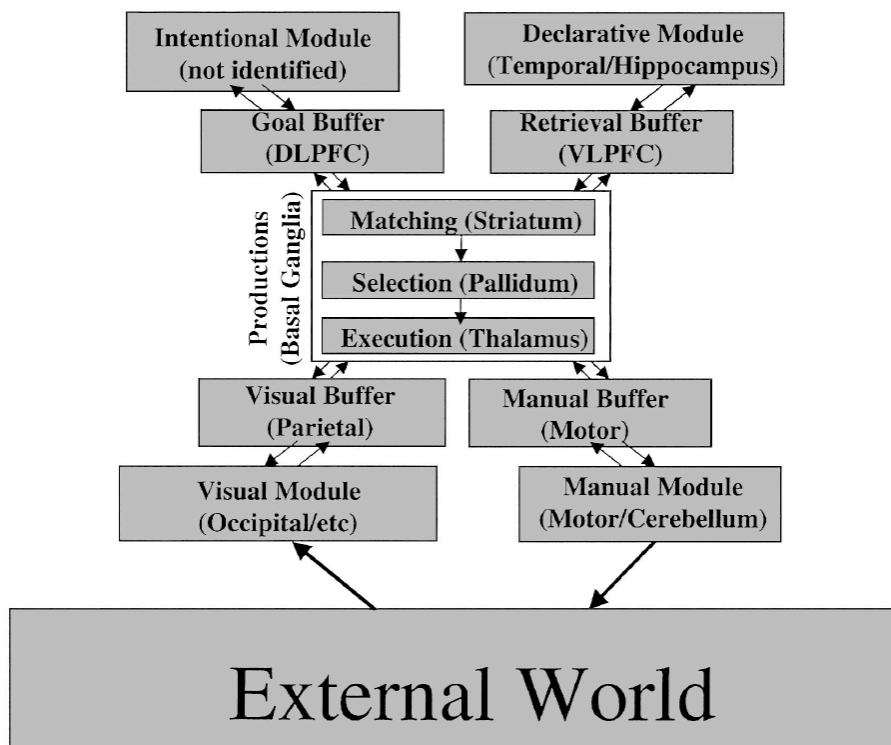


Figure 2: The information flow in ACT-R 5.0 (from [16]).

2.2.3 Conclusion on Cognitive Architectures

There are many similarities between these two main cognitive architectures. They are in agreement on symbolic representations at high levels of abstractions and they both represent operational knowledge by production rules. They both put emphasis on learning mechanisms based on remembering and abstracting previous successful courses of actions. They also both are not very much concerned with real time operation and building representations from sensory data. More precisely, they both say this is important, but at the same time they don't provide a clear approach for achieving it. The question of linking symbolic and subsymbolic representations is actually not really addressed. From a robotics standpoint, this however is a central question. We will draw some inspiration from those cognitive architectures, mainly in the way we can represent and apply operational knowledge.

2.3 Robot Control Architectures

2.3.1 Robot Control

The main kinds of activities usually going on within a robot system can be summarized as follows:

1. To interpret a mission or a task communicated by a user (e.g., reach some goal, grasp a given object).
2. To plan the necessary actions enabling to fulfill the task (action or motion planning).
3. To execute these actions, with adequate transitions, refinement and instantiation according to the state of the robot and of the environment (using sensing and actuation processes).
4. To take into account real-time constraints, and contingent events during action.

The robots developed in laboratories and institutions worldwide do not all exhibit these four activities. In some cases, the task is predefined, and there is no communication nor interpretation, and no planning. Some systems have poor adaptation to new events and lack reactivity. This is due to fundamental differences in the design of robot system architecture. The approaches differ in several manners, sometimes by the philosophical standpoint itself: some projects aim at imitating living beings, while others are more AI oriented, and others again have addressed real-time constraints and robots acting in real environments. The first trend of research sought inspiration from biology and ethology and has mainly yielded stimuli-response based systems. The second trend tried to use symbolic representations and some reasoning capacities, while the third has proposed a variety of architectures, more or less principled or ad hoc, and has finally come up with the present rather consensual "hybrid architectures" concept.

In order to better introduce the concepts underlying the architectural designs, it is useful to look at the concept of autonomy. We can distinguish two different kinds of autonomy:

Operational autonomy: defines the robot's ability to act (and react) in its local environment using sensing, motion, localization, obstacle avoidance, sensori-motor loops, etc. The operational autonomy depends on the performance of the perceptual processes, motion controllers, and their integration.

Decisional autonomy: defines the robot's ability to deliberate and to reason on its own tasks and to take decisions on the way they are ordered and achieved, and to supervise the execution of its actions. This usually requires abstract representations of the environment and automated reasoning capacities.

Fully teleoperated systems lack both kinds of autonomy. Usually operational autonomy is present in most systems to different extents. Behavior-based systems lack decisional autonomy in general, and are only able to perform reactively. Decisional autonomy requires complex abstract representations often difficult to build from sensory data, and time-consuming deliberation processes. Ideally, the control architecture must possess both kinds of autonomy and exhibit both decision-making and reactive capabilities, even if there are different approaches to achieve them. Indeed, the robot has to achieve tasks for which it is unavoidable to plan beforehand in order to decide for the best course of action, anticipating situations and expected events to monitor them. Tasks must be instantiated at execution time according to the actual context, because it is impossible in general to have a complete and precise knowledge of the execution situation. Finally, the robot must react in a timely fashion to events.

It is clear that the implementation of several task-oriented and event-oriented closed-loops for achieving both anticipation capacities and real time behavior cannot be done on a single system with homogeneous processes due to their different computational requirements. The architecture must therefore permit the integration of processes with different temporal properties and different representations. Therefore modularity and hierarchical design are two necessary

2.3.2 Examples of Robot Control Architectures

The design of robot control architecture is a long standing issue in robotics. We will not provide here a complete overview of control architectures, but rather name a few ones that represent major trends or approaches.

One general paradigm is based the "perceive-plan-act" cycle, which was criticized for its centralization and the lack of a reactive component. One example is TCA ("Task Control Architecture") [51], in which a central system decides the interactions among the components and resources allocation. The tasks belong to different categories (e.g., planning or execution or monitoring), and are scheduled for timely execution to cope with new events.

Brooks [21] proposed the "subsumption" architecture which falls on the other extreme and introduced a different approach known as "behavior based robotics". Here, the emphasis was more on the reactive capacity of the system and on its robustness. The organization is that of fixed layers of "sense-react" behaviors with growing complexity, all acting concurrently, the higher ones subsuming the lower ones, which are operational again when the higher levels are not operating (figure 3). Typically, the systems based on this paradigm do not include a deliberation component and are not taskable. Some authors [36] extended this paradigm by adding a deliberative component, which main role is to arbitrate among the behaviors and to coordinate them.

Hybrid architectures are those which include deliberative and reactive components together. There is a wide variety of such architectures developed concurrently over the last fifteen years at different research labs and a full overview is beyond the scope of this paper. They decompose the robot system into two or three main levels, having different temporal constraints and manipulating different data representations.

As an example, we can cite 3T [19] or Atlantis [29] which are three-layer architectures. They comprise a set of "skills" in one layer, sequenced by a component in another layer. The third layer is a planning system that reasons on goal achievement including timing constraints.

The Aura architecture proposed by Arkin [17], figure 4, initially focussed on navigation, integrates a deliberative component upon a reactive one. The control is based on the notion of schemas inspired from biology which represent a generic behavior associated with a perceptual input. The architecture was later extended to include learning and interaction capabilities.

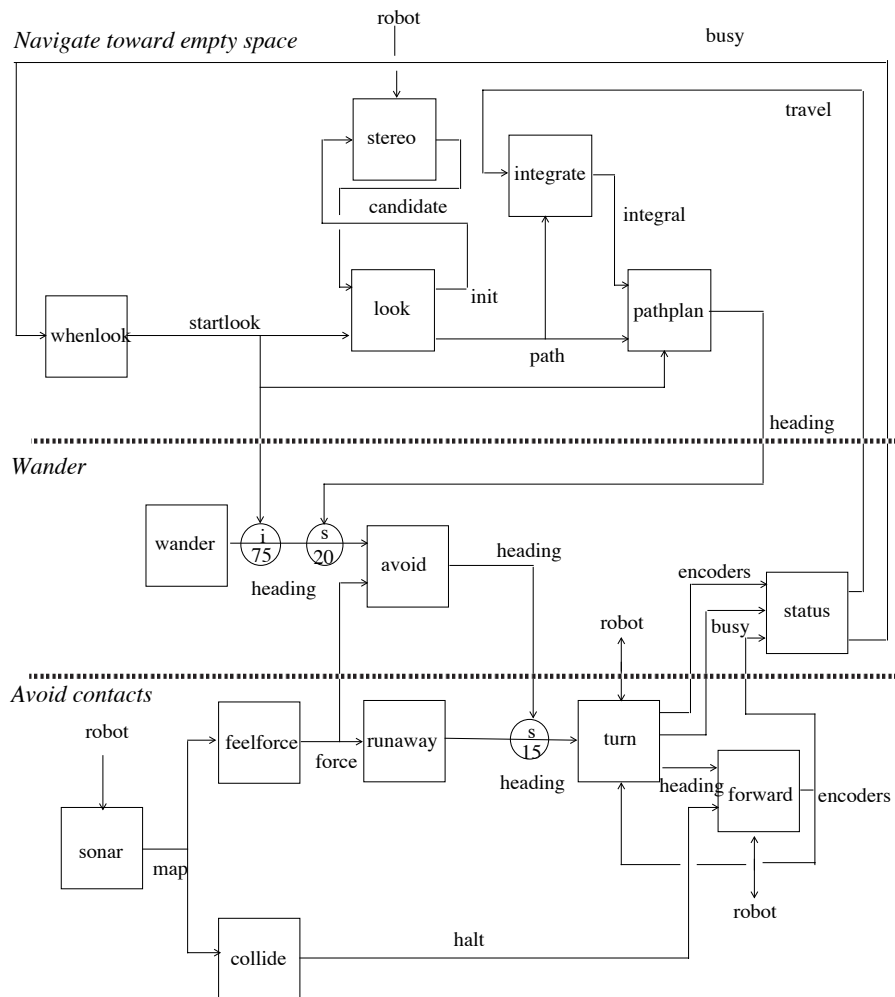


Figure 3: Instance of the subsumption architecture.

The LAAS architecture [15], figure 5 identifies three mail levels:

- a *functional* level: it includes all the basic built-in robot action and perception capacities. These processing functions and control loops (e.g., image processing, obstacle avoidance, motion control, etc.) are encapsulated into controllable communicating modules. In order to make this level as hardware independent as possible, and hence portable from a robot to another, it is interfaced with the sensors and effectors through a logical robot level. The modules are activated by the next level according to the task.
- a *decisional* level: this level includes the capacities of producing the task plan and supervising its execution, while being at the same time reactive to events from the previous level. The coexistence of these two features, a time-consuming planning process, and a time-bounded reactive execution process poses the key problem of their interaction and their integration to balance deliberation and reaction at the decisional level.
- At the interface between the decisional and the functional execution levels, lies an *execution*

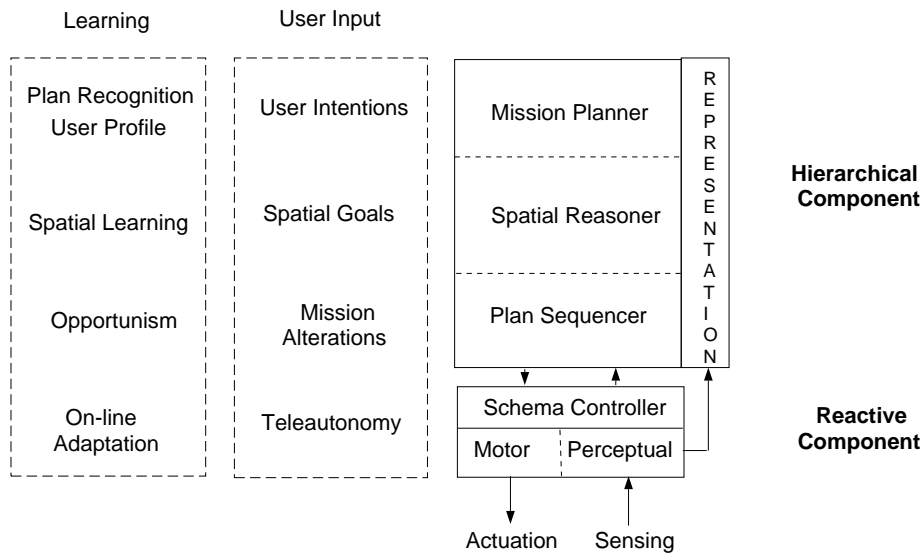


Figure 4: The general Aura architecture.

control level, which controls and coordinates the execution of the functions distributed, in the modules according to the task requirements. Despite its clear role in the architecture, it can actually be embedded in the functional level, given its intricate interaction with its components, resulting in a two-level architecture.

The *functional level* includes all the basic built-in robot action and perception capabilities, encapsulated into controllable communicating modules. These modules are activated by requests, send reports upon completion and export data. A module has the responsibility of the resources it manages (sensors, actuators) and acts upon requests by means of a set of specific processings (algorithms). This decentralized “responsibility” has two main objectives: (i) to simplify and to relieve the central control: the notion of service hides the complexity and the dynamics of the low levels processings, (ii) to increase the robustness of the whole system by introducing control and recovery procedures at the lower level.

The services are parameterized and activated asynchronously through a non-blocking client/server protocol: a relevant request, that may include input parameters, applies to every service of each module. Thus requests start processings. An ongoing processing is called an activity. The end of the service is marked by a reply returned to the client that includes an execution report and possibly data results. For instance, the execution of a trajectory is an activity that runs until the trajectory is over, or until an error occurs. In this last situation, the execution report characterized the failure (important drift, obstacle on the way, ...) that can be recovered by the decisional level. Every module of the functional level is an instance of a generic module which formal model is a finite state machine. A description language allows to define the module: all the services it manages, their input and output parameters, the associated code and their real-time characteristics (period, delay), the execution reports. From this description, the module is generated by a specific tool called GENOM (Generator of Modules) [27].

The *execution control level* filters the requests according to the current state of the system and a formal model of allowed and forbidden states. These rules and constraints are enforced in real-time and will

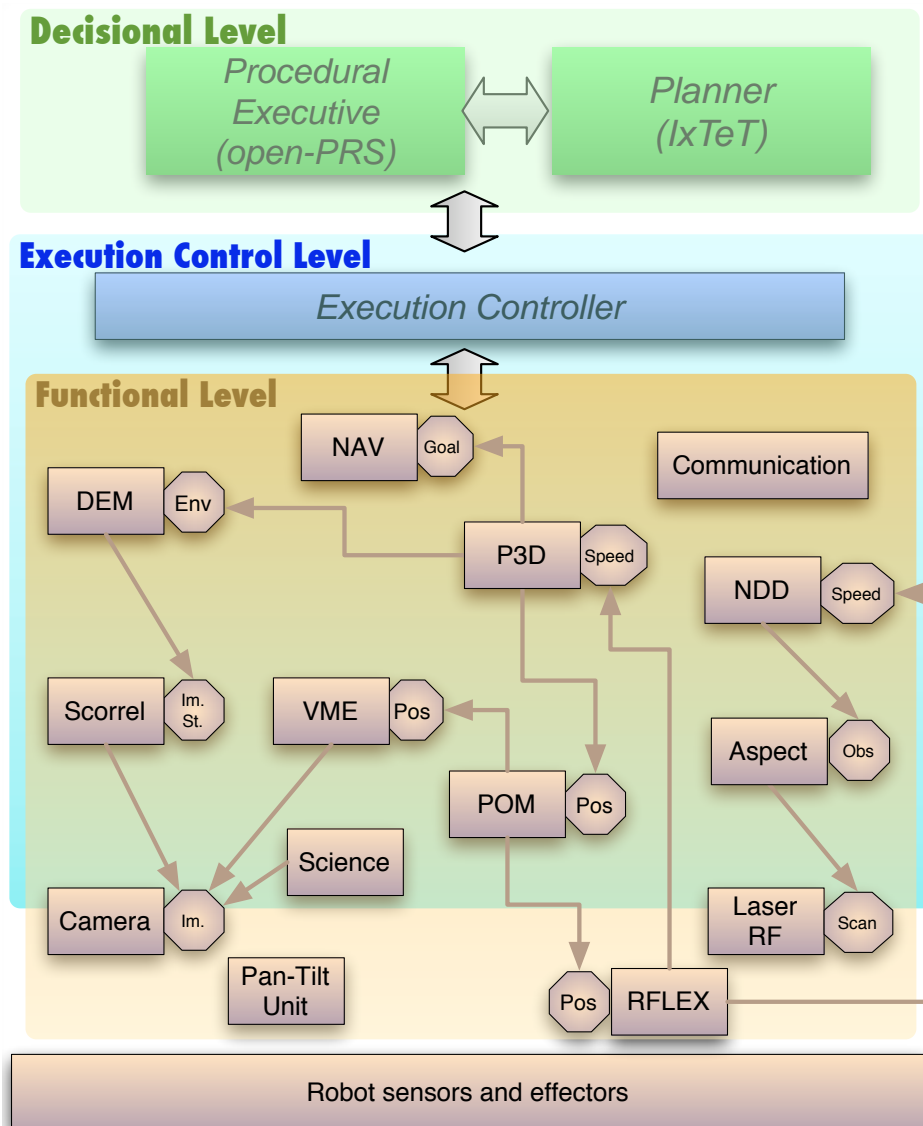


Figure 5: Instance of the LAAS architecture.

guarantee that the robots does not engage in any dangerous state [44].

The decisional level integrates deliberative planning, plan repair and execution control that takes into account resource level updates and temporal constraints. These processes are embedded thanks to two interacting components which explicitly represent and reason about time: a temporal planner and a temporal executive.

The plan execution is controlled by both procedural and temporal executives. The temporal executive decides when to start or stop an action in the plan and handles plan adaptations. The procedural executive (OpenPRS) expands and refines the action into commands to the functional level, monitors its execution and can recover from specific failures.

The CLARAty architecture [56], figure 6, proposes an organization of software components along two “axes”: an abstraction axis (i.e. from low level hardware to high level software processing) and

a functional/decisional axis. This results in an architecture which has two layers (a functional layer and a decisional layer). Yet, in each of these layer, one can have components at a very low level of abstraction (i.e. close to the hardware) from which one can build up higher level software components. For example a D* based navigation component could build on top of “grid map building” component, which itself would use a laser range finder, etc... Note that this hierarchy of functional components is not specific to CLARATy, the LAAS architecture above proposes a similar organization of its functional level. Similarly, on the decisional layer, one can have low level of abstraction decisional components (mostly executive components), up to high level goal planner (deliberative components). So the main difference of CLARATy over “standard” layered architecture lies in the capabilities of the system to perform “direct communication and control” between decisional components and functional ones of the same level of abstraction (i.e. without going down the level of abstraction of the functional layer). As a consequence, one could have a system which deliberates upon functional components of the lowest level of abstraction. This could be done by having decisional components directly interacting with the low level functional components. Another “strong” aspect of Claraty, is the use of object oriented components which ensures the proper interface with the components of the same layer and the one of the layer above.

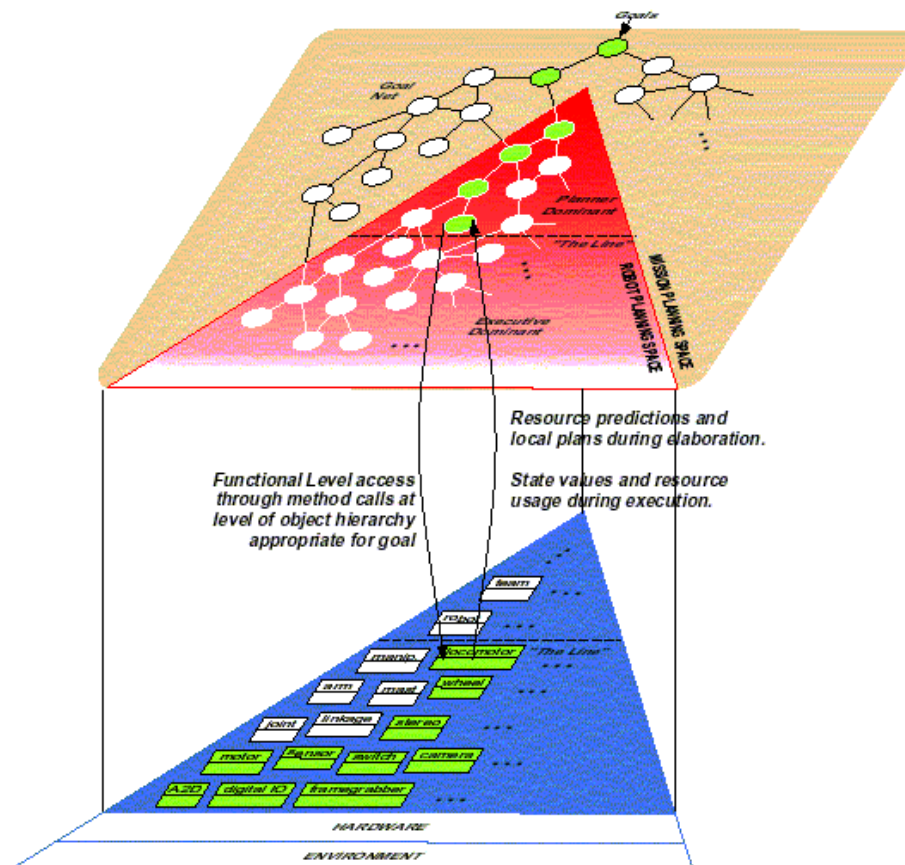


Figure 6: The CLARATY architecture.

2.3.3 Conclusion on Robot Control Architectures

These architectures are rather designed to control autonomous robots. They are preprogrammed and have little ability for learning. Humans interacting with the system (if any) are usually specialists, and this interaction is through a programming or command interface. In Cogniron, we are interested in an architecture that embeds "naturally" interaction with non-specialist users. There is nowadays a consensus on three-layered architectures integrating decision-making, execution control, and reactive sensori-motor loops (sometimes called behaviors), and we will adopt this structure for which we have a long experience for the decision/action part of our architecture. However, as we shall see, we will define an additional level on top of it to deal with multiple goals.

2.4 A Conceptual Architecture for Cogniron

2.4.1 Requirements and Architecture

We can classify a robot's activities into five main categories: perception, decision-making, action, communication, and learning. These activities run concurrently and permanently and use the set of basic functions the robot is endowed with. For example, the action of grasping an object is a sensori-motor process during which the object is perceived and modeled, decision on grasp position and robot trajectory is taken, and the grasp itself accomplished by the actuators, and the grasping skill can be improved and better learned.

The architecture that combines these activities should enable autonomy and adaptability: the robot should be able to carry out its actions and to refine or modify the task and its own behavior according to the current goal and execution context as perceived. The robot should also be reactive: it has to take into account events with time bounds compatible with the correct and efficient achievement of its goals (including its own safety). The robot's behavior should also be consistent: it has to be able to focus its attention and to react to events according to its objectives, not just according to events in a "stimulus-response" scheme.

Another desirable feature is robustness. This is not about adaptability for accomplishing the task, but more about the way the control architecture is able to exploit the variety of the sensory motor abilities and the redundancy of the processing functions to continue to operate even if there are changes in the environment and robot state. Robustness will require the control to be decentralized to some extent. Finally, extensibility, i.e, the capacity of integration of new functions should be possible. In classical robotics, this is achieved by adding modules for example, and we require the architecture to make this easy. But in our context this pertains to learning and "growing" new capabilities.

Finally, we want to study robots in interaction with people. Therefore, they should exhibit communication capacities, and be able to understand tasks that humans want them to accomplish.

In summary, the following requirements for a robotic system are to be considered:

- **Taskability:** The robot should be able to achieve multiple tasks described at an abstract level (language, gestures,...). Its architecture should therefore enable to easily and dynamically combine its basic functionalities according to the task to be executed.
- **Autonomy and adaptability:** the agent should be able to carry out its actions and to refine or modify the task and its own behavior according to the current goal and execution context as it perceives it.
- **Reactivity:** the agent has to take into account events with time bounds compatible with the correct and efficient achievement of its goals (including its own safety) and the dynamics of the

environment.

- **Consistent behavior:** the reactions of the robot to events must be guided by the objectives of its task.
- **Robustness:** the control architecture should be able to cope with failures, in particular by exploiting the redundancy and modularity of the processing functions and subsystems.
- **Extensibility and learning:** the architecture should provide for learning new knowledge and skills and integrate them. The global design should be modular to allow better extension.

These requirements call for the coexistence of both deliberative and reactive behaviors in the system, and for the integration of interacting subsystems performing according to different temporal properties.

The architecture proposed in figure 7 complies with these requirements.

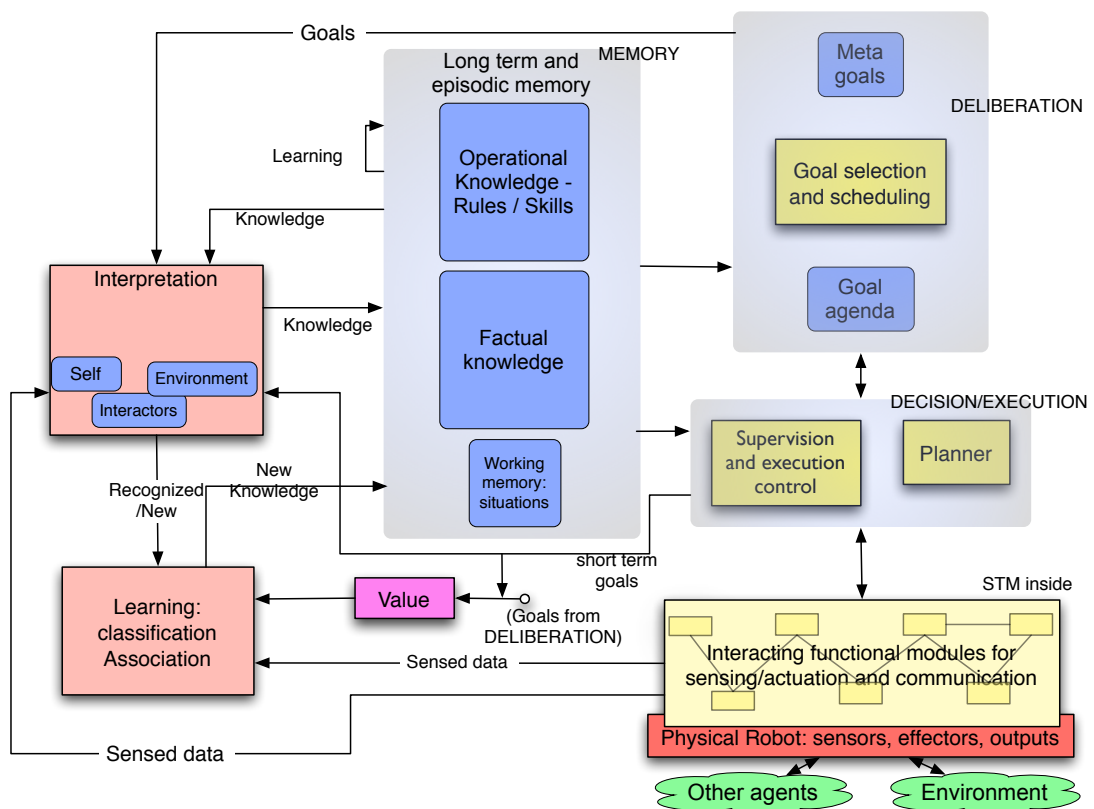


Figure 7: The proposed global conceptual architecture for COGNIRON. This architecture is a model of the concurrent processes and their interactions. A "box" in this figure is not necessarily a uniquely geographically located process; it rather represents several processes participating in the same function. Memory and representations are distributed across the processes.

This architecture is in some aspects similar to Aura, when considered at an abstract level (there are functionalities of learning and interaction), but the nature of the exchanges, the global economy of the

system is different. There are interaction and learning capabilities in SOAR and ACT-R as well, that we also consider differently here (see next).

2.4.2 Cognitive capacities

The Central and Pervasive Role of Learning. We contend that a cognitive robot must have, at the core of its operation, a permanent learning activity. This is indeed the only manner to achieve an open-ended system, i.e., a system able to operate without strong a priori limitations due to pre-defined representation primitives or operational capabilities. The cognitive architectures mentioned above, all account for a learning process which is however often implemented as an afterthought, or only in a specific manner (e.g., producing new rules from older ones).

We think that learning should be present at all levels and in every process. Indeed:

- There is a permanent flow of data incoming from the robot's sensors and low level processing functions. The perceived environment includes known and unknown objects and situations. Therefore a permanent process should be able to interpret the flow, discriminating what is already known, and classify data to produce new representations of space, objects and situations. The architecture has to include such an interpretation/learning activity. Interpretation and learning concern the environment, and the robot itself. In our context of a robot companion, we put also a special focus on the robots' interlocutors and interactors who are not just part of the environment. To cope with the information flow, not all incoming data should produce new categories. A saliency measure (or information measure) can filter out part of the data.
- The robot acquires knowledge from perception (including through interaction), but it also should be able to produce new knowledge through internal structuring, associations, deduction, abduction and induction mechanisms. The memory that contains factual knowledge is not just a repository of sensed data but an active process which transforms this data into linked structures through those mechanisms.
- Robot operational knowledge in terms of skills, rules and scripts (i.e. organized procedures) have to be permanently modified and improved, through their evaluation with respect to the achieved tasks in a supervised or a non-supervised manner. The chunking mechanism of SOAR can produce more successful new rules, but these are implicitly present in the system. Of course such a feature is desirable, but we seek a learning process that produces new skills from more basic ones, by associating perception and actuation to produce sensori-motor representations. This association must take into account the context and the pursued goals. We propose to do this by the use of a value function that expresses how much the associations enable to accomplish the goal, i.e., their utility, taking also into account their cost.
- This value function itself should not be predefined once and for all, but learned so that new associations can be made according to new goals of the robot. This can be a reinforcement learning process.
- The basic components of the system (sensor-motor loops, processing functions) have to be tuned to better execute. Local learning processes could be implemented at this level as well.

Perception and Interpretation. It has already been largely discussed in the design of control architectures that the system's knowledge cannot be stored in a central unique data base. Representations,

are rather distributed, and are adapted to the processes that handle them (or even confused with them in the case of neural architectures). The architecture we propose has this distributed memory feature. Each processing function has its own representations (e.g., the local environment representations for obstacle avoidance are points or edges for the avoidance servoing), whereas there is a more central representation of a situation as a result of the interpretation process. The coherence of all these representations is an important issue, addressed by Horswill in CEREBUS [31] where he proposes an approach based on a tagging scheme so that a given object can have different representations which remain related to each other.

The "interpretation" in our architecture is a data abstraction process that combines sensed data and extracts more global and semantic representations from them. The issue of coherence is to be addressed in the different instances of the architecture.

The "Memory" in the architecture is distributed and comprises factual knowledge, which are the representations the robot knows and their temporal relations (episodic memory), and operational knowledge which is the set of rules and procedures representing its capabilities and skills. A situation resulting from interpreting the perceived world in a working memory is included here.

Decision-Making and Reflection. We consider that the robot has permanent goals on the one hand (such as keeping its energy level sufficiently high, or keeping human satisfaction high, or keeping its integrity), called "metagoals", and goals that arouse from the actual situation as perceived by the interpretation process. In general the robot will be facing multiple conflicting goals, and even if some of the robot's goals are strong, to be implemented by sensori-motor reflexes (such as non collision), the robot might find itself in situations where other priorities have to be taken into account as well. The architecture includes two main decisional processes, named "deliberation" and "decision/execution". The role of the first process is to solve those multiple goal situations and produce a "goal agenda" that will be in turn solved by the decision/execution level whose role is to plan the course of actions that will accomplish these goals. In deciding on its goal agenda or in deciding the course of actions, the system uses the knowledge in the memory, including the operational knowledge and robot state.

Note that the decision/execution system is the decisional component of a two-level architecture, and the set of "interacting modules" is its functional level (see figure 5) (the "execution control level" is included in the supervisory module). Thus the conceptual architecture embeds a classical hybrid architecture. The planning system here is not in charge, but is used as a resource when necessary.

The set of sensori-motor functional modules operate and interact in a very similar manner to what we mentioned in the LAAS architecture for example, base on a Finite State Machine model. However, the detailed operation is not the topic here. This will be part of the actual implementation in the key experiments.

Interaction. Human-robot interaction and communication is not explicitly depicted in the architecture schema in the figure. The processes responsible for interaction are part of the perceptual, decisional and action processes. One can consider that there is an interaction process "floating" upon the architectural components that are part of an interaction (data processing, language understanding, dialogue generation, etc.). Interaction is therefore an intricate activity of the system itself. The inputs from robot sensors include the communication inputs and the actions include those that are produced for communication. Interpretation of the exchanges with humans is done within the general interpretation process, and shown in the "interactors" box. Dialogue and other interactions are part of the decisional processes. The interaction processes are detailed in the second part of this report.

2.5 Future Work

The conceptual architecture above is not in its final design. There are a number of open issues that have to be addressed in the next phase. Each process will be studied in more detail, as well as their interactions, the data they exchange, and their temporal constraints. Data coherence across the representation is also a key pending issue.

The design of the architecture corresponds to the requirements listed above. It is taskable because of the interactions capabilities, and it is able to accomplish goals autonomously, to anticipate future events, and to reflect on the robot's own capabilities because of its deliberation and planning components which use a model of those capabilities (the operational knowledge). It is reactive (it embeds a hybrid architecture). The modular and decentralized interacting processes can be redundant, and there can be several different means to accomplish a given goal (different skills) that the system can use. Finally it is extensible because of the learning component.

The instantiation of the conceptual architecture and its implementation in the key experiments is the scientific method to prove its viability, its ability to correspond to the requirements listed above and to provide for the sought cognitive capabilities. We have started to map the conceptual architecture with the architectures implemented in the key experiments (see RA7 deliverable in which we map the conceptual architecture to the experiments). Validation through the evaluation of those experiments is under way, and will be an important effort in the next phase.

3 A Framework for Human-Robot Collaborative Problem-Solving

This section focuses on the management of interaction with humans as an integral part of the robot control architecture.

Indeed, we aim to come up with a principled way to deal with Human-Robot Interaction (HRI) for robot task achievement in presence of humans or in synergy with humans.

We have devised a decisional framework for human-robot interactive task achievement [23]. We briefly present here below.

Besides, together and in coherence with this framework, we intend to develop and experiment various task planners and interaction schemes that will allow the robot to select and perform its tasks while taking into account explicitly the human abilities as well as the constraints imposed by the presence of humans, their needs and preferences. In addition to the constraints that have been studied last year, we have investigated issues linked to the strong relationship between geometric constraints and symbolic task and behaviour constraints.

3.1 Context

In our context the human is physically present in the vicinity of the robot, is sensed by the robot and may even participate to the task performance.

In relation with this, a number of recent contributions about close interaction deal with the notion of physical and mental safety [41] or the introduction of emotions and/or cognitive models in robotic structures [20, 38]. Very often, HRI is merged into the task performance. This tends to reduce HRI to a (sometimes very sophisticated) human interface.

Our aim is to endow the robot with an explicit consideration of humans and with the ability to manage its interactions with them. This must be considered at different levels: at the architecture level as well as at the task/motion planning and execution level.

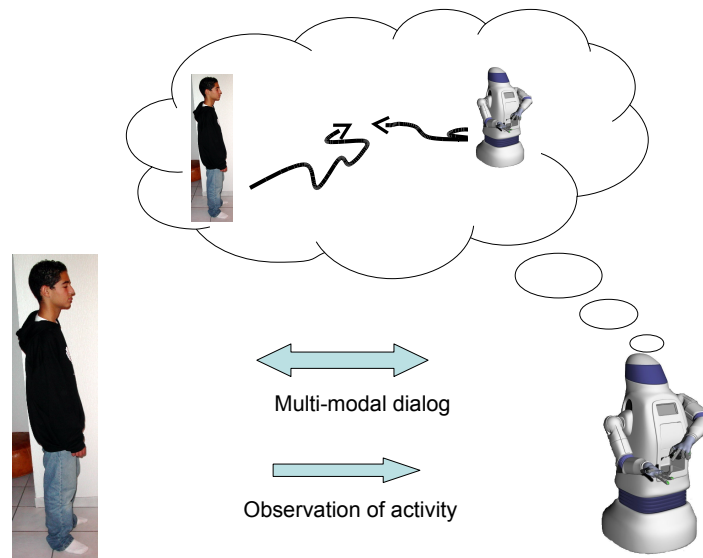


Figure 8: Reasoning about HRI and anticipation of human activities: sources of information are multi-modal dialogue, and observation of environment and human activity

It is worth noting that there are several architectures that explicitly embed interaction [52, 32, 49, 28]. One key source of inspiration is the Joint Intention theory [25, 24, 33]. It is based on the notion of commitment for team members and defines for a team the concept of Joint Persistent Goal. These definitions constitute a basis for the elaboration of cooperation schemes between heterogeneous agents. We follow a stream similar to [26, 22, 54]. Indeed, we believe that an effective implementation of this theory can be done, when limited to a clearly defined context in which the robot will deal explicitly with the actions, beliefs or intentions of the human partner.

3.2 A Decisional framework

Our robot is controlled by a three layer architecture [15]. We present briefly the design of the decisional layer in which we have introduced what we call InterAction Agents (IAAs). They are similar to proxies but are directly implemented on the robot side as a representative of a human agent. To make the interaction more explicit we have defined a complete process of establishing a common goal, achieving it and verifying commitment of all agents involved. Besides, relevant IAA models should be devised and used in the robot planning activities. Such models will range from high-level specifications of the human abilities and preferences to geometric attributes such as position, posture or field of view regions.

We envision HRI in a context where two agents (a human and a robot) share a common space and exchange information through various modalities[23, 1]. Interaction happens as a consequence of an explicit request of the human to satisfy a goal or because the robot finds itself in a situation where it is useful if not mandatory. In both cases, the robot has a goal to satisfy. An important issue is the notion of engagement, a process in which the robot will have to establish, maintain and terminate a connection with a human partner. Besides conversation, such a process will provide a framework for robots performing tasks in a human context. This covers goal establishment, selection of an incremental refinement of the task that is intended to be achieved, and execution monitoring. This

context will be used by the robot in order to follow human task performance, to monitor his/her commitment to the common goal, and even to influence it.

The proposed decisional framework [23] consists of several entities, having each a specific role as illustrated by Figure 9. The HRI we consider in this context is the common achievement of tasks by two agents - a robot and a human - in order to satisfy a joint goal. The human involvement may range from a direct participation to the task achievement, to a simple “acceptance” of robot activity in his close vicinity.

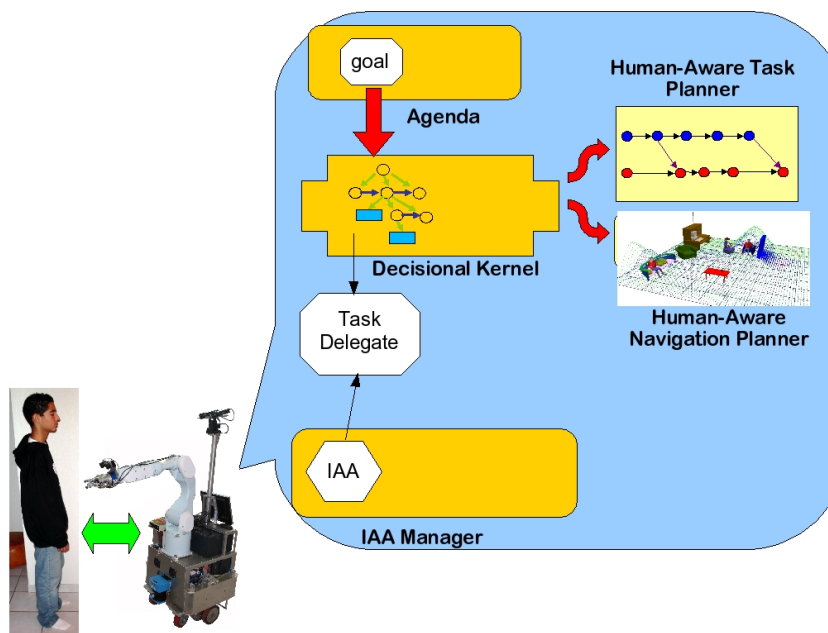


Figure 9: Decisional framework for a HRI-enabled robot

The Agenda. Several goals may be sought at a given time, involving possibly several persons. At any moment, there may be several active, inactive and suspended goals. The Agenda manages the current set of robot goals. It ensures the consistency between active goals, and determines their priorities, and their causal links. Based on data provided by the Supervision Kernel, the Agenda determines the relevance of goals and decides to create, suspend, resume or abandon a goal. When a goal is created, it may be associated to the robot alone or to a “team” of agents.

The IAA Manager. The humans encountered by the robot are represented by entities called “Inter-Action Agents” (IAAs). An IAA is created dynamically and maintained by the “IAA Manager”.

The Task Delegates The set of active goals entails the incremental execution of a set of tasks in interaction with humans. Each task corresponding to an active or a suspended goal is represented by an entity called “Task Delegate” that is in charge of monitoring the progress towards the goals of both the robot and the IAA and to assess the level of commitment of the associated person.

The Robot Supervision Kernel The Robot Supervision Kernel is responsible of all tasks selection, refinement and execution. It maintains an integrated view of all robot activities and ensures a global

coherence of robot behavior. It is the only entity that can send execution requests to the functional level.

For each new active goal, the Robot Supervision Kernel creates a Task Delegate, selects or elaborates a plan and allocates the roles of each team member.

For all the other active goals, the Robot Supervision Kernel has already a plan and is in charge of the execution of the robot part. Whenever an elementary action is performed, the Robot Supervision Kernel forwards this information to all active Tasks Delegates. Depending on the context, the planning process can be more or less elaborated. The planning activity associated to a task is a “continuous process”; it provides, incrementally, the next sub-tasks to achieve. It has also to state, depending on the context, on the feasibility or relevance of the task.

3.3 Human-Aware Task Planning

Context. The main point here is how high level robot task planning skills should be developed in order to allow it to act as an assistant. The robot must not only perform its tasks but also act in a way judged as “acceptable” and “legible” by humans. In such a scheme, the robot plans for itself and anticipates the human behavior in order:

- to assess the feasibility of the task (at a certain level) before performing it
- to share the load between itself and the human
- and to explain/illustrate a possible course of actions.

Representing social constraints. We have studied in the past period a planner that is able to take into account “social constraints” and to synthesize plans compatible with human preferences, acceptable by humans and easily legible in terms of intention. It is based on HTN (Hierarchical Task Network) planner SHOP2[39] mainly because it permits to specify costs for actions and encode procedural knowledge. Both the robot and the human are represented in terms of actions they can perform. A “team” composed of two “agents” (the robot and a human) can be represented as: $(A_{human}, C_{human}^{ctx})$ and $(A_{robot}, C_{robot}^{ctx})$ where A_i are sets of actions and C_i^{ctx} are their context-dependent associated costs.

The introduction of costs allows to select preferred behaviours. Indeed, at this level, it is possible to deal with social constraints that can be represented as:

- costs/utilities that denote the difficulty and the pleasure an agent has in an action realization
- undesired states (from the human side)
- desired or undesired sequences of actions that may induce a robot behavior that is not understandable (legible) by its human partner
- synchronizations and protocols that may represent social or cultural conventions

Examples involved domestic like situations where the robot essentially performs fetch-and-carry and cleaning tasks in interaction with a human. This study has confirmed the relevance of this level and of the types of considerations that should be taken into account when building robot plans in this context. Results and analysis of user studies have showed that people comfort is influenced by the behaviour of the robot. It has also been shown that the context plays an important role as well as the spatial relationships such as the Human to Robot (H-R- approach preferences. This has a direct consequence

on high-level task refinement which should be ideally based on a coherent formalization that allows to take into account various constraints of different nature. For instance, some of them can be best expressed geometrically while others may be expressed in terms of temporal or causal links between actions.

This is why we have intended to apply and extend models and algorithms similar to those we have developed in aSyMov, a planner that is able to deal with intricate symbolic and geometric constraints [30].

The example illustrated in Figure 10, is extracted from a plan produced by aSyMov. It represents two snapshots from a plan that integrates not only geometric and symbolic constraints of a “classical” pick-and-place task but also H-R approach motion preferences (approaching from the front, waiting agreement before entering intimate human spatial zone).

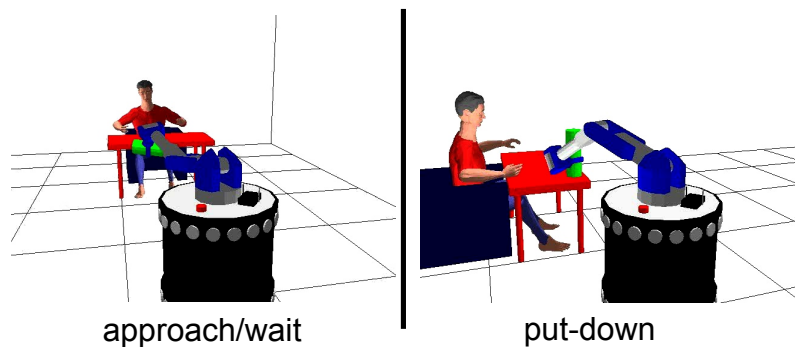


Figure 10: Planning a fetch-and-carry task for a person

As a consequence, we orient our developments towards a task reasoning system that is able to answer various questions involving symbolic (in the broad sense) and geometric constraints: placement of robots and objects, visibility criteria relative to the human and to the objects, spatial placements required by the dialog, spatio-temporal synchronization, ordering and synchronisation of human and robot actions, etc. While the general scheme we propose might be difficult to implement in a general context, we believe that it is a reasonable challenge to implement it in the case of a personal robot assistant essentially devoted to fetch-and-carry, as well interactive manipulation tasks and navigation activities, i.e. the contexts envisaged in KE1 and KE2.

4 Models and Algorithms for cognitive robots supervision

We report here on the design and implementation of a robot supervision system based on joint intention theory that is intended to provide a framework for human-robot interaction at planning and execution level and for incremental context-based task refinement and execution.

4.1 Design considerations

We briefly summarize here the joint intention theory (JIT) [35, 25] and the necessary adaptation to its use in our field. Several contributions such as TeamWork[48, 49, 50, 52], Retsina [28, 42] and Polyscheme [53] have tackled similar problems including human in the loop. Our system is closer to [26] where the robot selects a target for interaction based on the perceived desire of human subjects to interact with it using JIT via communicative acts.

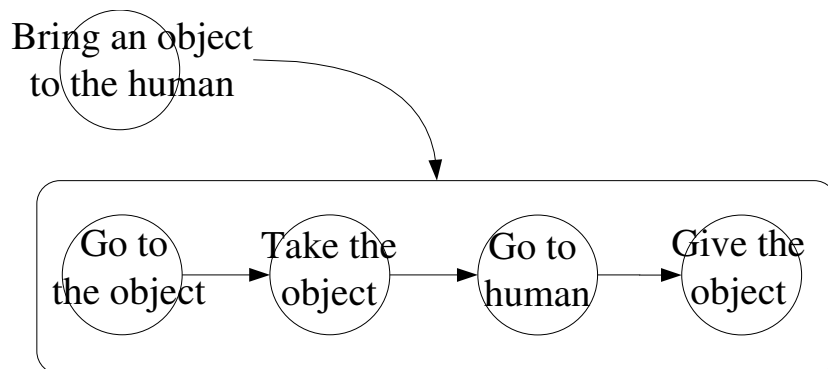


Figure 11: Decomposition of a task: 'Bring an object to a person'

Next we illustrate the individual case and joint case of commitment and intention, and joint commitment establishment and end on an example.

Example

Let us consider a “simple” task where the robot must *Bring an object to the human*. Primarily we decompose the task in an ordered list of sub-tasks: *go to the object*, *Take the object*, *go to the place where the human is*, *Give the object to the human*. This view, illustrated by Figure 11, just takes into account basically at a given level of granularity what the robot must do to achieve the task.

Individual Case

In Cohen and Levesque [35, 25] theory of commitment, the basic definitions are:

- Individual Commitment: An agent has a persistent goal relative to q to achieve p in the case:
 1. she believes that p is currently false;
 2. she wants p to be true eventually;
 3. it is true (and she knows it) that 2. will continue to hold until she comes to believe either that p is true, or that it will never be true, or that q is false.
- Individual Intention: An agent intends relative to some condition to do an action in case she has a persistent goal (relative to that condition) of having done the action and, moreover, having done it, believing throughout that she is doing it.

These definitions stress the difference between a commitment and an intention. In the first case, the agent has to make sure that p will be achieved. In the second case, the agent will achieve p . So, one can be committed to an action and never have the intention to do it.

Robot behaviour: In our case, p represents a task for the robot e.g. *Bring an object to the human*. Figure 11 represents the way the robot intends to achieve the task. When it chooses this plan, the robot commits for each sub-task as the task develops. When committed to a task, the robot can't give up unless one of the defined conditions is satisfied. If an intention to achieve the task fails, the robot must retry.

When committed to a task, the robot must be able to check if p is currently false (the task is unachieved), if p is true (the task is achieved), if p will never be true (the task is impossible) or if q is false (the task is irrelevant). In JIT, q represents the irrelevance clause that allows to end a task without any other reason.

The robot must be able to follow the progress of the task *Bring an object to the human* and to have abilities to know/monitor when it is unachieved, achieved, irrelevant or impossible. While achieving this task, it must also do and follow the progress of the current sub-task e.g. *Take the object* and need to know/monitor (by various means) when it is unachieved, achieved, irrelevant or impossible. Hence, the robot must be able to do that at different levels. In Polyscheme [53], “Focus of attention” determines to which aspects of the world the system will devote its representational and inferential abilities. The need to monitor task progress can be seen as a focus of robot’s attention concerning its tasks.

The level of granularity will allow us to be more accurate. Indeed, the reason that makes the task, *Bring an object to the human*, impossible will not necessarily be the same that the reason that makes the sub-task, *Take the object*, impossible and the recovery action will not be the same neither. If the task *Bring an object to the human* becomes irrelevant, all the sub-tasks that the robot intends to achieve will also become irrelevant. However, if *Take the object* becomes irrelevant, only this sub-task becomes irrelevant, the task *Bring an object to the human* has to be done and so, the robot must find another way to intend it, for example using re-planning capabilities.

At this step, we see the difference between commitment and intention and the necessity for a hierarchical representation. The next subsection considers the integration of the human in the loop.

Joint Commitment

“without a shared mental state a team does not exist” [25]

Here again, we start with some definitions:

- Joint Commitment [25, 35]: Two agents x and y have a joint persistent goal p with respect to q when precisely the following conditions hold: there is a mutual belief that p is not currently true, it is a mutual goal to bring about p , and p will remain a weak mutual goal at least until there is a mutual belief that p is either true, or will never be true, or the relativism condition q is no longer true.
- Weak Mutual Goal: A weak mutual goal (WMG) is a mutual belief that each agent has a weak achievement goal (WAG) towards the other for achieving p relative to q .
- Weak Achievement Goal: An agent x has a WAG towards another agent y when the following holds: if agent x believes that p is not currently true then it will have a goal to achieve p , and if it believes p to be either true, or to be impossible, or if it believes the relativism condition q to be false, then it will have a goal to bring about the corresponding mutual belief with agent y .

In our case, p represents a joint task. Mutual belief represents a belief that is shared by the two agents, i.e. one agent has a belief about something AND has the belief that the other agent has the same belief AND vice versa. This notion of mutual belief, “shared mental state”, is central in JIT. This will have design consequences on the robot abilities to reason about task performance in human environments. These definitions and more generally joint intention theory have been successfully used in multi-robot cooperation or in human-robot cooperation [48, 49, 50, 52]. However, the most studied context involved a human and a remote robot system. The problem, in such a context, is essentially to find

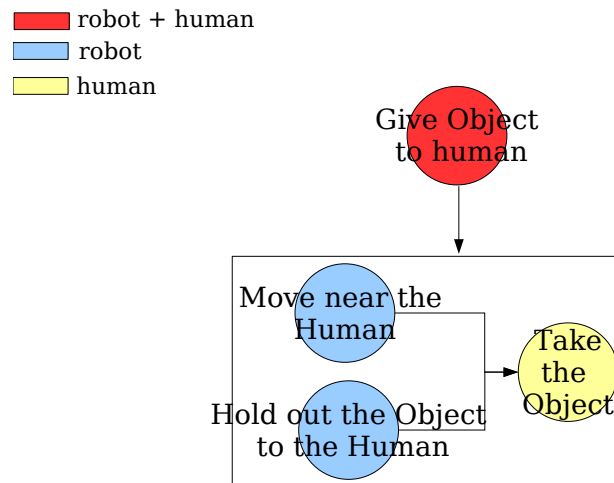


Figure 12: Decomposition of a “simple” task ‘Give an object to a person’ into sub-tasks involving the robot, the person or both.

the right level of autonomy for the robot depending on the context. As it is explained in [37], they deal with *adjustable autonomy*, i.e. “agents dynamically varying their own autonomy, transferring decision making control to other entities (typically human users) in key situations”. The question here is “determining whether and when transfers of control should occur from the agent to other entities”. In our case, the question is slightly different, determining whether and when the robot should interact and/or take initiative towards the human, whereas the robot and the human try to achieve a common task. The human and the robot share a same environment, they are often close to each other and perceive each other’s activity. The question is to equip the robot with suitable context-dependent abilities to achieve its tasks in the vicinity and/or in interaction with the human. We can call such issues “adjustable interaction”.

Robot behaviour: The task *Give an object to the human* can be considered as a joint task because it involves the robot and the concerned human (see Figure 12).

Following the discussion above, a joint task cannot be considered as achieved if, privately, an agent comes to that conclusion. The joint task continues until mutual belief concerning the task state is obtained. For example, in *Give an object to the human*, if the robot has the belief that the task is achieved, it has to alert the human (unless it already knows that the human knows the task is achieved), the robot and the human must reach a mutual belief about the task state. This means that the robot has the belief that the task is achieved AND the belief that the human believes the task is achieved AND the human has the belief that the task is achieved AND the belief that the robot believes the task is achieved.

Now, we cannot assume or impose to the human to follow a “strict” Joint Intention Theory protocol and, for instance, to alert the robot whenever he has the belief that the task is achieved. The robot will have to act, observe and infer in order to maintain a “shared mental state” that will allow it to achieve its interactive tasks.

Our system will be completely on the robot, it is the robot that will make all observations and deductions concerning the beliefs and commitments of the human (and it is inherently a partially observable information from the robot point of view when it is...) and to behave in a way that is easily understood by its human partner.

This has to do with perspective taking ([43],[53]). From [43], “perspective taking is the ability of a person to take someone else’s perspective”. We can assume that if the robot has to deal with its beliefs about human beliefs it should take his perspective. This can be interpreted at a “geometric” level (robot, human and objects relative placements) as well as “symbolic” level (goals, task achievement process, task and environment state).

We will see later that these beliefs will concern the commitment to the task and, once involved, the task state (achieved, unachieved, irrelevant or impossible).

Establish Joint Commitment

Now that we have defined commitment, intention and mutual belief. We would have a look on the way to establish joint commitment.

- Persistent Weak Achievement Goal (PWAG) [33]: An agent x has a PWAG toward another agent y when the following holds: if agent x believes that p is not currently true then it will have a persistent goal to achieve p , and if it believes p to be either true, or to be impossible, or if it believes the relativism condition q to be false, then it will have a persistent goal to bring about the corresponding mutual belief with agent y .
- Establishment of a joint commitment [33]: Mutual belief in each other’s PWAG toward the other to achieve a goal p is sufficient to establish a joint commitment to achieve p provided that (1) there is mutual belief that p has not already been achieved, and (2) the PWAGs are interlocking i.e., one PWAG is relative to the other.
- Joint Intention [25, 35]: A team of agents jointly intends, relative to some escape condition, to do an action iff the members have a joint persistent goal relative to that condition of their having done the action and, moreover, having done it mutually believing throughout that they were doing it, i.e. each agent individually intends his or her own action and is committed to the other’s action.

These definitions stress the fact that mutual belief in each other PWAG is sufficient to establish a joint commitment. The question is: how do we obtain and maintain mutual belief?

Many systems consider explicit communication: verbal [25, 35, 33] or software [49, 48, 52] is the mean to obtain mutual belief. These systems consider that all agents are willing communicators and sincere in that they never try to get other agents to believe something that they do not want them to know (though there were adaptations to deal with communication problems [48]).

Concerning human-robot interaction, on the one hand we cannot be sure that the human is a willing communicator, and on the other hand, if so, communication could be non-verbal such as a gesture, movement, etc., and difficult to interpret by the robot even if needed. However, whatever the form the communication takes, we will now refer to it as *communicative acts* as defined by [33]: “Communicative acts must be characterised as attempts because there is a possibility that the act may not succeed. For example, suppose that I sincerely request you to open the door. The goal of my request is that you open the door and the intention of my request is that it be mutually believed that I want you to open the door. My request is successful if you recognise that I want you to open the door and my request is satisfied if you actually open the door in response to my request. The best I can do is to make my intention known to you and it is up to you whether or not you actually open the door. If I have reason to believe that you have not properly understood my intention then I may repeat my request, i.e., I

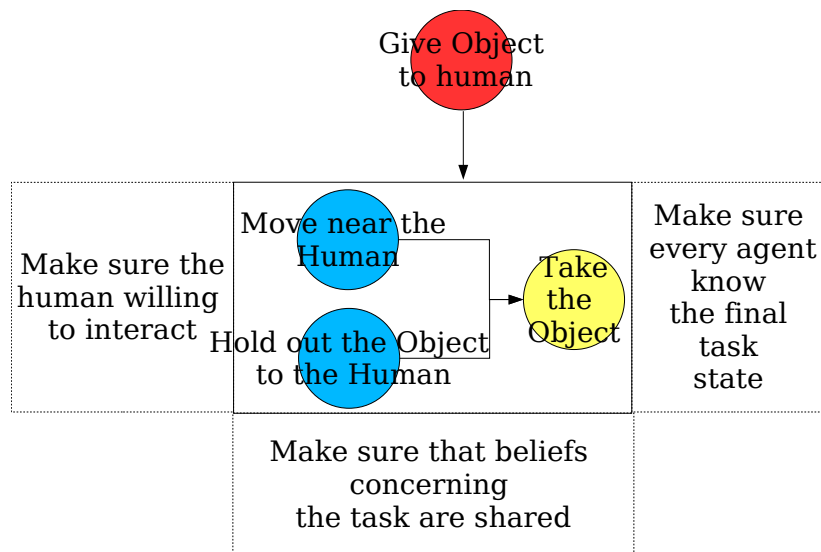


Figure 13: a simple task 'Give an object to a person': adding activities related to joint human-robot task achievement

may attempt again to make my intention known to you. Accordingly, attempt is defined as having a goal and an intention.”

In our system, we consider commitment establishment as a particular phase, called PRETASK, preceding the execution of the task itself. In this phase, the robot first informs the human about the joint task so the human and the robot share the belief that the robot will form a goal with the human. If it seems that there are communication problems during this information, the robot must retry.

On the other side, when a task is ended, we have seen that communication should take place to inform everyone of the task state. This communicative act aims to gain mutual belief concerning the task state. We also consider this as a particular phase that we call POSTTASK.

Now, during the execution of the task, it could happen that the robot got uncertainty concerning the human involvement in the task, i.e it could happen that the robot believes that the human believes the task is irrelevant. In that case, we have defined a special phase, named CHECKTASK where the robot goes to check its information e.g. asking oral confirmation to confirm a fact, or behaving in order to explicitly check the human “motivation”.

Finally, we call the (effective) execution phase DOTASK. During this phase the robot follows the progress of the beliefs concerning the task. Figure 13 and 14 show a joint task development and needed monitoring.

4.2 Supervision system

Based on this analysis [1, 2], we have developed and implemented a supervision system. This first implementation has been realized not only as an endeavour to investigate human-robot interactive task achievement but also to establish a principled framework for testing the first KEs scenarios. The supervisor is constructed in order to perform incremental context-based task refinement and execution based on protocols derived from joint intention theory.

Each human entering in the field of the robot is taken into account as an agent with whom the robot might collaborate or at least interact (these agents are named IAA for InterAction Agent in [1, 2]).

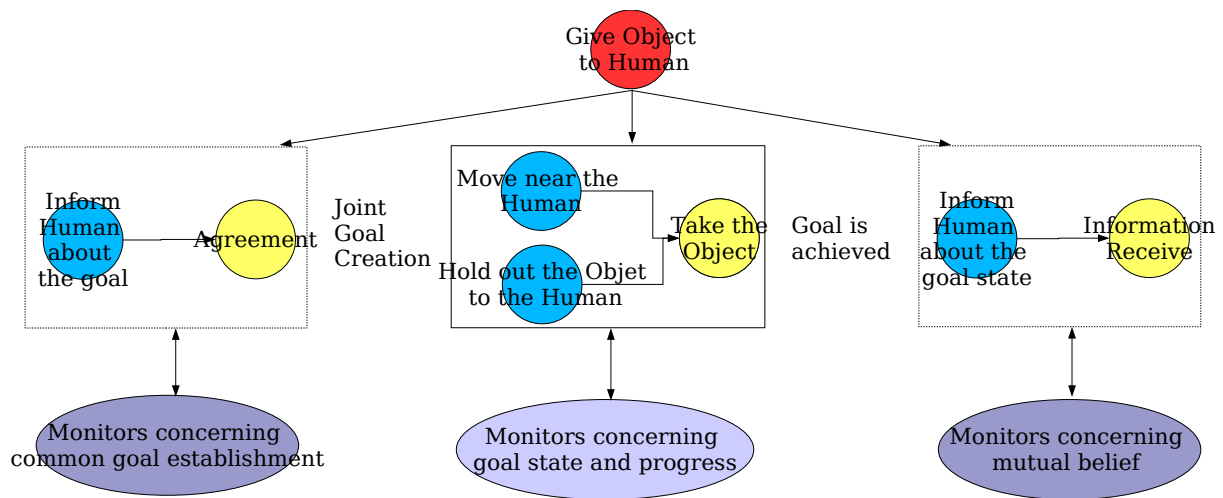


Figure 14: a joint task 'Give an object to a person' with associated monitoring activities

To be homogeneous, the robot itself is also considered as an agent. For now, agent will refer to the human or to the robot, unless the word "robot" is used explicitly.

First, we briefly present next the State Space that is maintained, the different kinds of tasks that have been defined and the task refinement and execution procedure.

State Space

It is mainly composed of:

- Agents Commitment to the task (= AgentCo (example: RobotCo)): Committed, Uncommitted (default).
- Agents Belief concerning the task state (= AgentTS (example: RobotTS)): Unachieved (default), Achieved, Irrelevant, Impossible.
- Robot progress in the task (= RobotP): Idle, Start, Planning, In Task, Stopping, Stopped.

Agent Commitment and Belief concerning the task exists for the robot only in the individual case, and for the robot and other agent(s) involved in the joint task case (one by agent). Robot progress in the task corresponds to the state the robot is within the task.

Task definition

Considering the difference between the individual and the joint case, and the granularity available within task/sub-task hierarchy, we have defined four task types:

- **Root Task:** in the current implementation, the goal agenda is limited to one top level robot task,
- **Individual Task:** a task that can be decomposed but that concerns only the robot,
- **Joint Task:** a task that can be decomposed but that concerns the robot and another agent,
- **Activity:** a non-decomposable task that corresponds to a low-level routine for the robot, or an atomic human activity as observed by the robot.

```
(=>(PLAN (ROOT)
  (. (. Robot (INITIALISATION) .)
     (. Robot (LOOKING FOR INTERACTION)
              (. Robot Agent (DOSERVICE) .) .))
  .))

(=>(PLAN (GO_TO $destination)
  (. (. Robot (FIND TRAJECTORY $destination) .)
     (. Robot (MOVE $destination).)
     .) .))
```

Figure 15: Examples of Task Plans stored in the plan library

A task is defined by:

- its current plan: corresponding to the set of activities and/or sub-task(s) the robot will do to achieve the task, plan examples are shown in Figure 15,
- a set of monitors that will follow the progress of the task and verify whether the task is: still unachieved, achieved, irrelevant or impossible. An example of monitors definition is given in Figure 16.

In the current version of the system, the monitors are passive i.e. a monitor cannot be the cause of a robot action. They just use information provided by the task performance to conclude. An active version of these monitors would correspond to the Observers defined in [1, 2] but are not yet implemented. Monitors are a convenient way to deal with multimodality as it is possible to associate an arbitrary number of monitors corresponding to different modalities.

Context-based task refinement is performed through a (re)planning step. In the current version, plans are hand-coded and stored in a plan library (Figure 15). However, provision is made to connect various types of planners. Figure 17 illustrates the dynamic task structure and planning contexts.

The planning activity associated to a task is a “continuous process”; it verifies periodically the feasibility of the task and provides the next sub-tasks to achieve, depending on the context. Events detected by monitors at a given level are propagated through the task structure and cause sub-task creation, destruction, (re)-planning.

Joint Task Each Joint Task follows a common framework. The plan of a Joint Task is divided in four sub-tasks (Figure 18):

- PRETASK: corresponding to the establish commitment phase,
- DOTASK: corresponding to the execution of the task,
- POSTTASK: corresponding to the end of the task,
- CHECKTASK: corresponding to the treatment of potential problem concerning mutual belief between the robot and the agent involved.

```
(=>(MONITOR (GO TO $destination) $who1 $who1 ACHIEVED OK (GOTOACHIEVED)))

; watch BRAKESON
(=>(MONITOR (GO TO $destination) $who1 $who1 IRRELEVANT BRAKES (RFLEXTRACKSPEEDSTARTBRAKESON)))

; watch ASPECTFAILED
(=>(MONITOR (GO TO $destination) $who1 $who1 IMPOSSIBLE ASPECT_FAILED (ASPECTFROMPOSTERCONFIGFAILED)))

; watch people blocking the robot
; few time
(=>(MONITOR (GO TO $destination) $who1 $who1 IMPOSSIBLE PEOPLE (ATBLOCKED)))
(=>(MONITOR (GO TO $destination) $who1 $who1 ACHIEVED NEXT_TO (ATNEXTTO)))

; delay
(=>(MONITOR (GO TO $destination) $who1 $who1 IMPOSSIBLE DELAY (ELAPSED-TIME (TIME) 600)))

; segloc
(=>(MONITOR (GO TO $destination) $who1 $who1 IMPOSSIBLE LOCLOOP_FAILED (SEGLOCLOCLOOPFAILED)))
```

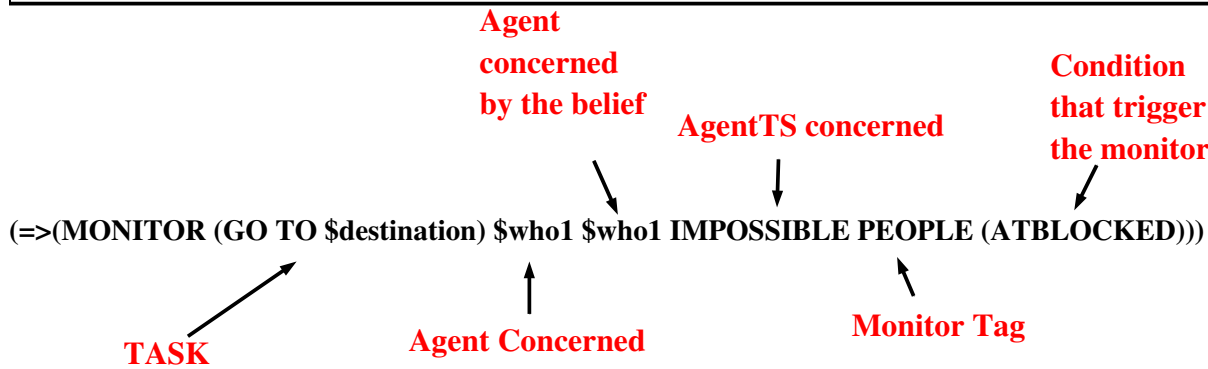


Figure 16: Syntax for encoding the set of monitors associated to the task (GO TO \$dest)

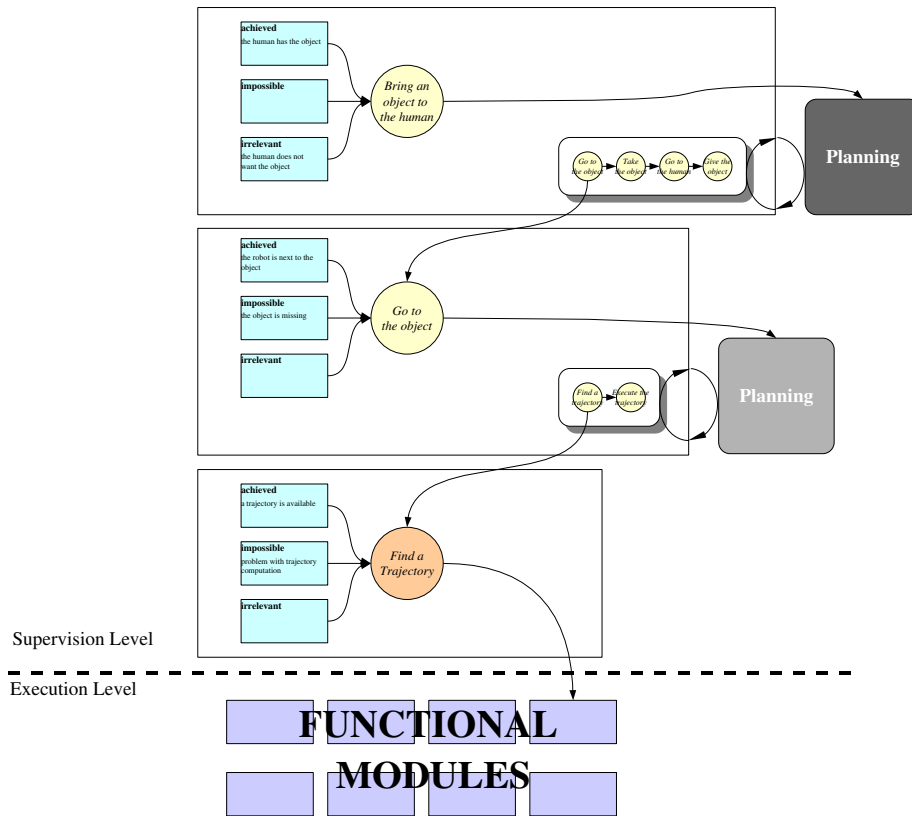


Figure 17: The dynamic hierarchical task structure managed by the supervisor

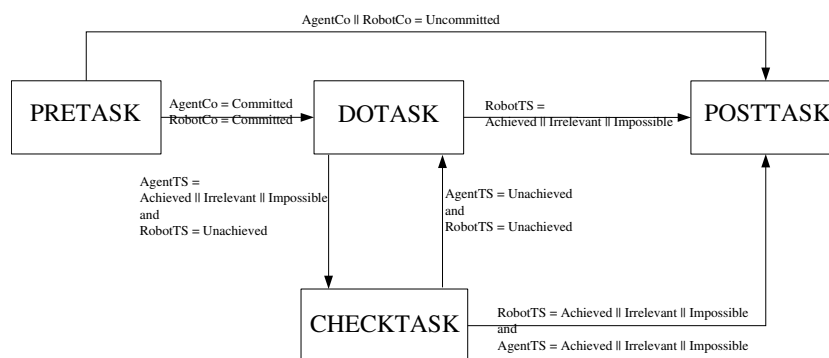


Figure 18: Joint tasks embed four sub-tasks: (1) PRETASK: to establish commitment phase, (2) DOTASK: effective execution of the task, (3) CHECKTASK: treatment of potential problem concerning mutual belief between the robot and the agent involved. (4) POSTTASK: corresponding to the end of the task.

Each item above can be defined as an individual task, a joint task or an activity.

Concerning Joint Task, several kinds of monitors could be defined. In figure 16 showing the individual task (GO TO \$dest), we see that monitors are defined for \$who1 concerning \$who1, i.e. monitors involving the belief of the robot concerning itself. In the joint task, we may also define monitors for an agent concerning another agent. We define, for example, monitors for an agent concerning the robot. Notations are the same for all task types.

4.3 Results

We have developed and implemented a robot supervision system that deals with tasks in terms of individual tasks, joint tasks and activities. Each task is defined by a plan and dedicated monitors. A plan corresponds to a sequence of sub-tasks and/or activities. Monitors serve to state whether a task is unachieved, achieved, impossible, irrelevant or stopped.

The system can be controlled at different levels at the same time. If something is detected at a given level, the system is able to take it into account at that level by applying adapted solutions and propagating, when necessary, events towards the higher or the lower levels.

In the current implementation, the supervisor is written in Open-PRS. The task plans are hand-coded (a set of pre-defined task library) and only the robot is able to propose a task. However, the supervisor is designed to take into account future extensions involving on-line task planning.

The system has been tested on Rackham (see RA7 - KE2 deliverable for a description of the robot) during its last visits at the Space City Museum in Toulouse ([23]). The root plan of the robot was: *Initialisation*, *Looking-for-a-visitor* and *Guide-visitor*. Initialisation was only done once, then the robot switched between the last two tasks.

During four days, Rackham ran in average one hour per day (the only time limitation was due to the mandatory presence of a Space City staff member for security reasons). 25% of the time the robot has run the task *Looking-for-a-visitor* and 75% of the time, Rackham run *Guide-visitor*.

Moreover, we have used the same system to integrate and run tests involving Human-Detection and Tracking Services that have been developed in RA2. In that case, the plan was: *Initialisation*, *Wait-For-People*, *Want-to-interact?*, *Guide-visitor*.

Wait-For-People and *Want-to-interact?* are two complementary tasks. *Wait-For-People* detects by the help of the Motion Monitoring if someone enters in the field of the robot. If it happens, *Want-to-interact?* checks if the person is interested by the robot i.e. approaches the robot (Body Tracking, Face Tracking and finally Face Detection) and/or touches the touch-screen, etc. If nothing happens for a while in this second task, the robot returns to *Wait-For-People*. Otherwise, *Do-Service* is launched. *Do-Service* is a version of *Do-Guide* that uses all the services provided by Human-Detection and Tracking Services whereas *Do-Guide* uses only face detection.

Compared to the previous supervision system, it is easier to add new capabilities to the robot and to program new tasks involving interaction. In addition, monitoring is facilitated by the help of layered levels that are all reactive.

4.4 Future work

We have implemented a first version of the supervisor based on the design proposed in[1, 2]. The results so far are encouraging, and we believe that is a good basis to implement and investigate human-robot interactive task achievement.

The next objectives are:

- the integration of human aware task and motion planning

- the extension of the agent models used by the supervision
- the possibility to deal simultaneously with several agents-tasks pairs
- the investigation of the links with multi-modal dialog
- further investigation on high-level goal agenda and its link with robot initiative.

5 Advances in Robot Control Architecture Implementation

Work has been done on the development of the software architecture and tools of the robots that are used in the key experiments.

5.1 BIRON architecture

At the end of the second phase the BIRON architecture [10], is based on a hybrid control mechanism and has three layers: a reactive, an intermediate and a deliberative layer. Modules that are responsible for reactive feedback of the system are set on the reactive layer: the *Person Attention System* detects potential communication partners and the *Object Attention System* combined with the information of the Dialog provides a learning process for new objects [11](see also RA1, RA2) that users refer to. Since these are purely data-driven processes they belong to the reactive layer. Modules responsible for higher-level processing that involve top-down, expectation-driven strategies such as a planner or the Dialog System for interpreting user utterances based on the current system state, are located on the deliberative layer.

The system is centrally controlled by the so-called *Execution Supervisor* on the intermediate layer. It coordinates the module operations and makes sure that neither the reactive layer modules control the deliberative layer modules nor vice-versa. Instead, it exerts control by taking into account the overall system state. This is contrary to most hybrid architectures where a deliberator continuously generates plans and a reactive plan execution mechanism just has to make sure a plan is executed until a new plan is received. As a consequence this architecture allows for both fast reaction to dynamic environmental changes and extensive high-level planning and reasoning activities.

In order to satisfy additional system safety requirements, modules should fail perceptibly as in a real world robot failures cannot be excluded. We realized this feature by messages which are initiated by the main loop and sent in fixed intervals to modules being in communication with the sending module. If a module does not receive these messages anymore, it can determine that the corresponding sender stopped working correctly. In this case corrective actions can be taken to recover from the failure. If recovery is not possible then the robot is at least able to ask the user to call technical support using the dialog system which is also responsible for handling miscommunication.

For combining data of the mentioned different modalities a first step towards an *Active Memory* representing several aspects of the scene based on the processed data has already been made by memorizing data over time about persons interacting with the robot which are supplied by the person attention system.

5.2 Control Architecture tools related to “Programming by Demonstration”

UKA has pursued its development of a task representation for service robots which is set up of parametrised basic actions. These actions are linked within the task description in a tree like manner, and the tree is read and executed in a depth- first search. The task representation is able to cope

with exchange of variables, and program parts can be reused in other situations and contexts. The tree-like structure also supports this re-usability. This approach of modeling task knowledge [4] has proven to work on a real robot, giving the possibility to automatically choose the best alternative solution for each sub-task at run-time. The design of the task knowledge representation further allows for task adaptation and extension at run-time. This will be the next step in the development of Flexible Programs.

IPA has dedicated effort to the means by which the components of a cognitive architecture, once defined, can work together. The goal is to provide a framework/toolbox, so that the various contributions from Cogniron can fit together within a Key Experiment or beyond. Therefore tools were developed allowing a dynamic and fast adoption to requirements defined and being defined in RA6. Cooperation took place between several partners on this subject, especially UH and UKA.

5.2.1 Developments

IPA's high level robot control language "Go" [18] was extensively enhanced. "Go" was so far used for controlling lower level architectures of robots. With the help of several developments and ongoing improvements we want to create a tool for implementing robot architectures. "Go" is based on the script language Python [45] and is implemented as module within. Go has been released as free software under the GPL.

To the already existing functionality the following components were added.

XML interpreter for learned actions: The work of UKA concerning the "Programming by Demonstration" part of KE3 recognizes actions made by humans and records them in an XML format. The learned actions are stored as a hierarchical tree of XML macros. We developed an XML interpreter for processing the generated XML macros containing learned actions and executing them on any robot using "Go".

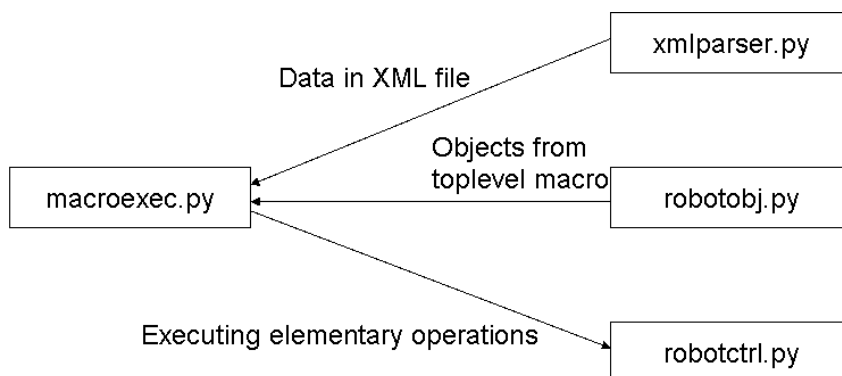


Figure 19: Modules of the XML interpreter

Four modules were developed (Figure 19) [55]. One module (xmlparser.py) parses the XML macro using a SAX parser [47] and returns all the information relevant for the XML interpreter. The module macroexec.py gets the output and determines the execution order of the task macros (sequential or parallel). Two modules (robotctrl.py, robotobj.py) describe the actual robot hardware and environment and have to be written for the specific a set-up.

Data management using Smart Pointers Due to the vast amount of data delivered by modern sensor (High resolution 2D colour cameras, 3D TOF cameras, etc) some methods need to be developed for data exchange. This is especially necessary, if several components share the same data of one sensor. Smart pointers [55] were developed, allowing different code sections the access the same object. The advantage of the smart pointers developed here are the possibility to share these pointers between C/C++ and Python, i.e. “Go”. Focus was also laid on the possibility to exchange these smart pointers over a network for distributed computing, thread safety and automated reliably garbage collection.

Others UH has been using Python for the control of their robots. To make the integration of “Go” possible for them, an extensive user manual was written for “Go”. UH is currently working on the integration of “Go” and is planning to write some of their own extensions (APIs and GUIs for their specific hardware).

For testing and debug simple GUI functionalities were implemented into “Go”.

Development components were made available to other partners via the Cogniron wiki.

5.2.2 Future work

The next steps will involve:

- Analysis of data flow requirements between CFs: Design and implementation of architectural tools addressing real time data flow between components (eg. direct connections between components for fast and data intensive communications controlled by “Go”). Development of real-time design patterns for service robots.
- Visualization of robot architecture structures extracted from a “Go” framework. Implementations of tools for configuration, development, verification and benchmarking of architectures.
- Design and implementation of a framework and tools for HRI components, based on results and experiences of HCI.

5.3 LAAS Open Software for Autonomous Systems

LAAS has worked on an open version of GeNom, called OpenGeNOM, which has been released on October (URL: <http://softs.laas.fr/openrobots/>)

The main tools that have been released are:

- GenoM, a set of tools for constructing and encapsulating elementary robot functions.
- OpenPRS is an open source version of PRS (Procedural Reasoning Systems) / Propice.

Besides source-code and installation procedures, scientific papers, as well users manual are available at the specified URL.

Open Genom is used in KE2. It has also been installed on KE1 platform. OpenPRS has been used to encode the current version of the KE2 supervisor.

LAAS Open Software for Autonomous Systems



These pages contain the collection of software that was developed at LAAS/CNRS for the *Architecture for Autonomous Systems* project. All together, the software provided here tackle the study and design of autonomous machines integrating perception, action and reasoning capabilities.

As of now, we have released only a very little subset of what we would ideally see here. There is a great difference between a software used internally and one that is publicly available: the latter must be sufficiently documented and easy to use so that users give it at least a try. This effort requires time. So as time will go on, you might see more and more interesting software appearing...

Latest News [\[more\]](#)

- 18 Nov 2005 - [pocolibs 2.1](#)
- 04 Oct 2005 - [libedit 2.10 and eltcsh 1.6](#)
- 29 Aug 2005 - [mkdep 2.4](#)
- 26 Aug 2005 - [pocolibs 2.1beta1](#)

General Information

- [About](#)
- [People](#)
- [CVS Repository](#)
- [Report or query a bug](#)
- [Intranet](#)

Tools

- [mkdep](#)
- [eltcsh](#)
- [gdhe](#)
- [OpenPRS](#)
- [GenoM](#)

Libraries

- [editline](#)
- [pocolibs](#)

Web site: mallet@laas.fr
Last modified: Tue Nov 15 09:3:07 CET 2005

Figure 20: The set of tools currently available at “<http://softs.laas.fr/openrobots/>”

6 Intentionality Attribution

6.1 Overview

The central objectives of WP6.4 are to identify parameters that are important in how people attribute intentionality to robots. This report will provide a summary of the research activities carried out during 2005 to fulfil the outlined objectives of WP6.4, provide documentation of dissemination in relation to peer reviewed scientific papers and conference activities, and will provide an outline of future research objectives in relation to WP6.4 from M25-M42/48. Substantial time has been spent during 2005 on dissemination activities in relation to WP6.4, including further analysis and publication of results reported in the 2004 deliverable D6.3.1. This is illustrated in the annex of publications provided at the end of this document. As there are a number of publications, a summary of the most important findings will be provided and the reader will be directed to the relevant publication in relation to that particular research if further particulars are required.

To outline the research undertaken to meet the objectives of WP6.4, the remainder of this summary document is comprised of the following:

- Further analysis and publication of results from the 2004 studies regarding subjects' perceptions of robots and the relationships subject/robot personalities
- Further analysis and publication of results from the 2004 studies regarding subject's comfort-spatial zones during human-robot interactions
- Findings regarding people's comfort levels during various human-robot interaction scenarios
- New user study on how human subjects perceive different robot cues (e.g. gestures, light, sound) and different robot appearances, assessing subjects' as well as robots' personalities in a scenario where the robot needs to attract a person's attention
- Findings regarding how a robot can attract a person's attention. Future research objectives from M25-M42/48 for WP6.4

6.2 In depth evaluation of the relationship between subject and robot personality and subjects

Personality characteristics are used to assist people in the interpretation and understanding of different social situations and behaviours. Research studies have found that people tend to assign personality attributes to computers, agents, and robots, which could assist the user in understanding its behaviour by shaping people's expectations about the interaction experience. Therefore, personality characteristics could play a central role to understand the attribution of intentionality to robot behaviour. A central goal of this research was to consider whether humans try to match and project their own personality attributes and styles to that of a robot to create an engaging interaction that they feel at ease with and can make sense of. The alternative view is that humans may not want to perceive themselves as being similar to a robot in terms of personality attributes. If this is the case they may infer different personality traits, or no personality traits, in the fear of losing their own identity, and wanting to remain unrelated to a robot. The key questions addressed in this research were:

1. Are there significant differences between participant personality traits and assigned robot personality traits?

2. Is there a relationship between human and robot personality?
3. If humans do project their own personality onto robots, does this attribution depend on the way the robot behaves?
4. What are the design implications for robots based on the findings from personality theory?

In a user study carried out in summer 2004, 28 adults interacted individually with a non-humanoid robot that demonstrated two robot behaviour styles (socially interactive, socially ignorant) in a simulated living room situation, relevant specifically to KE2. After the interaction session, subjects completed a series of questionnaires that assessed the extent to which adult ratings of their own personality traits were similar or different to the two robot behaviours. During the first six months of 2005 the data was analyzed in greater detail (first results were presented in D6.3.1) as part of the second implementation plan. A main summary of the results indicated that:

- For individual personality traits, subjects perceived themselves as having stronger personality characteristics compared to the socially ignorant and socially interactive robot behaviour styles.
- Both the socially interactive and socially ignorant robot behaviour styles were perceived as more emotionally stable, more introverted and less psychotic than subjects' own personality.
- Overall, subjects did not view their own personality as similar to the socially interactive or socially ignorant robot behaviour style.
- Factors such as subject gender, age and technological experience were important in how subjects viewed their personality as being similar to the robot personality.
- The attribution of personality analysis revealed that subjects evaluated the robot as being more similar to themselves with respect to the extra-introvert factor compared to the neuroticism-emotional stability, and psychoticism factors.

The current results have important future design implications as they suggest that the intentionality attribution assigned by people to robot behaviour styles could be very different depending on age, gender and level of technological experience. These individual differences also have important implications for the overall engagement and satisfaction for different human-robot interaction scenarios. The above results have been analysed in more detail resulting in the submission of a paper [13]. Further analysis revealed that, as our trials used a revised and shortened version of the Eysenck PEN model of personality, we were not able to replicate the same personality factor types (i.e. extraversion vs. introversion, neuroticism vs. emotional stability, and psychoticism). Therefore, to explore the research questions of whether there were differences or relationships between subject and robot behaviour styles, factor analysis was carried out on the 13 personality traits that were identified as being useful for the study. Table 1 outlines the personality factors that emerged from subjects' ratings of their own personality characteristics, and subject ratings of the personality characteristics of the socially interactive and socially ignorant behaviour styles. This summary table provides some further tentative confirmation that subjects viewed their own personality as very distinct to that of the two robot behaviour styles. However, it also provides some suggestive evidence that subjects were able to attribute differences in personality styles according to the two pre-defined robot behaviour styles. For example the most important factor attributed to the socially ignorant behaviour style was termed "hostility" and the second most important factor for the socially interactive behaviour style was termed "outgoing and proactive".

	Subjects	Socially Ignorant Robot Behaviour	Socially Interactive Robot Behaviour
Factor 1	Social Confidence	Hostility	Emotionally Withdrawn
Factor 2	Sensation-Seeking	Emotionally Withdrawn	Outgoing and proactive
Factor 3	Spontaneity	Outgoingness	Authoritarianism
Factor 4	Strong-mindedness	Independence	Solitariness
Factor 5	Insecurity	/	/

Table 1: Summary of the personality factors for subjects, Socially Ignorant robot behaviour, and Socially Interactive robot behaviour

6.3 The association between peoples' personality and spatial-comfort zones during human-robot interaction trials

A further research area related to human-robot intentionality and behavioural attribution concerns spatial-comfort zones between humans and robots. Psychology research findings have demonstrated that social spaces substantially reflect and influence social relationships and attitudes of people. The user study carried out in summer 2004 investigated whether human-human personal spatial zones transfer to human-robot interaction. Two main hypotheses were investigated: 1) we predicted that approach distances preferred by humans when interacting with a robot would be comparable to those preferred when humans interact socially with each other, 2) we wanted to determine whether common personality factors exist, which could be used to predict subjects' likely approach distance preferences. Twenty-eight adults who interacted individually with a non-humanoid robot were asked during the trials in summer 2004 to approach the robot as closely as they felt comfortable to do so, then approach the robot as closely as possible, then to withdraw to a comfortable distance. This was performed twice, and each distance recorded. A similar test was performed with the robot moving towards a stationary human. Preliminary results on social spaces that subjects adopted with regards to the robot were documented in D3.1.1. The relationship to personality factors, relevant specifically to WP6.4, was investigated in more depth during the first 6 months of 2005 as part of the second implementation plan.

A summary of the main results indicated that:

Human to Robot (H-R) approach preferences and tolerance:

- Approximately 60% of subjects approached the robot to a distance consistent with the normal Personal or Social zones for human-human conversation.
- A relatively large number (40%) approached the robot at a distance consistent with the Intimate or Close Intimate on the Hall Proximity Scale.
- In follow up questionnaires, these 40% of subjects did not want to be intimate friends with the robot, so the subjects probably perceived the robot as an object or machine, rather than as a social being.
- The more "proactive" a person was, the greater the human-to-robot approach distance tended to be.

Robot to Human (R-H) approach preferences and tolerance:



Figure 21: Comfort Level Device to allow subjects to indicate their comfort levels during HRI trials.

- Most adult subjects (when standing and not moving) allowed the robot (moving towards the subject in a straight line) to approach to distances corresponding to the personal and social spatial zones and similar to the results above, with 40% of subjects allowing the robot to approach them right up to the 0.5m robot safety limit.
- The overall mean comfortable approach distance was 0.9m.
- A large majority (more than 80%) also indicated that the robot did not make them feel uncomfortable or threatened when approached closely (to the 0.5m safety limit) by the robot.

Paper [6] documents the results in more detail in the appendix.

6.4 Human comfort levels during human-robot interaction trials

A further aspect that our team has investigated is human comfort levels assigned to various robot behaviours during human-robot interactions. The advantage of these studies has been that rather than relying purely on questionnaire responses, a Comfort Level Device (CLD) was developed to allow subjects to indicate their comfort levels during various human-robot interaction scenarios (cf. D3.1.1). This novel method allowed us to consider human attributions and preferences for a variety of robot behaviours relevant to WP6.4. The handheld CLD (illustrated in Figure 21) enabled people to use a slide button to indicate their internal comfort levels, ranging from “very uncomfortable” to “very comfortable” throughout human-robot sessions. Each of the HRI sessions were videotaped and detailed analysis has been carried out to determine aspects of human-robot interaction behaviours that human users attribute as being comfortable or uncomfortable.

A summary of the main results indicated that: When humans and robot are moving within the same shared environment, humans were uncomfortable when the robot:

- moved behind them,
- blocked their path,
- moved on a collision path with the subject,
- moved around for no apparent reason,
- also, humans did not like the robot to interrupt them during tasks which did not involve the robot.

Papers [8, 9] document the approach and the results in more detail. Please refer to the appendices.

	Mechanistic Appearance	Basic Appearance	Humanoid Appearance
Type of Gesture/Cues			
Gesture Manipulator	Lifting gripper	Pointing arm	Lifting arm
Sound Method	Beep	Mechanical voice	Human voice
Light	1 light	2 lights for eyes	Face with eyes/flashing lights

Table 2: The different robot permutations designed for the study

6.5 How should a robot attract a person’s attention? Preferences toward different robot appearances, gestures and cues (new user studies in 2005)

An important consideration for WP6.4 is to consider people’s perceptions and preference attributions toward robot intentionality. To investigate this we have designed a study, and carried out user trials in November/December 2005 that consider peoples’ preferences towards different robot appearances, gestures and cues in a scenario where the robot needs to attract a person’s attention (a scenario defined to be specifically relevant to KE2). Table 2 outlines the different robot permutations designed for the study to evaluate whether people have different preferences for three robot overall appearance styles, for the task of attracting their attention. In addition to exploring people’s preferences toward different robot appearance styles, we were also interested in whether people attributed different personality characteristics depending on the appearance styles, and whether there were trends between subject personality types and overall robot appearance preferences. The specific research questions investigated were:

1. Did a common pattern emerge towards preferences for any of the robot appearance styles?
2. Did people distinguish between the different robot appearance styles in terms of personality characteristics?
3. Are there any observable trends between different subject personality types and preferences for different robot appearance styles?

Figure 22 illustrates the different robot configurations.

The scenario designed for these particular trials took place in a “real” home (The University of Hertfordshire ROBOT HOUSE) to increase believability and ecological validity of the trials. During the trials in summer 2004 many subjects commented on the “simulated” nature of the experimental environment (a lecture room transformed into a “living room”). The new experimental environment is an apartment rented in a domestic area in Hatfield, UK, furnished with “everyday items” used by tenants. Live trials (relevant to WP3.1) took place in the ROBOT HOUSE where also the scenarios for the cues/appearance trials were developed. Subjects were shown videos of the scenarios (cf. [12] for a justification of this methodological approach). To ensure standardisation for the different robot appearance styles, cues and gestures, the same scenario was used for the different robot appearance permutations. The scenario involves a person sitting comfortably at home reading a book and listening to music when the door bell rings. As the person is listening to music, he does not hear the door bell, so the robot tries to attract the person’s attention to inform them that someone is at the door. Each of the three scenarios to illustrating the three robot appearance styles was videotaped. This method

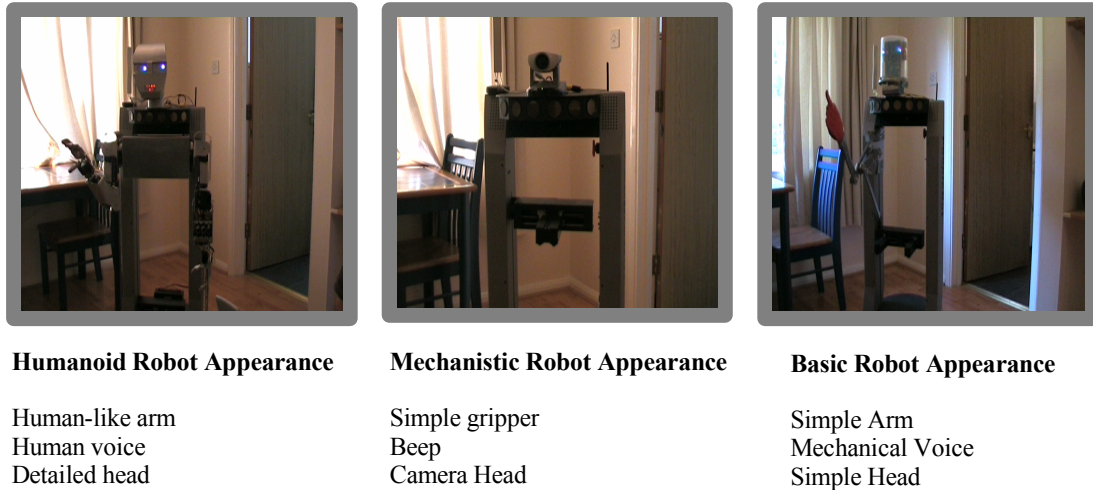


Figure 22: Robot configurations

was selected for a number of reasons, including the need to ensure standardisation for each of the appearance styles, it would have been extremely complicated to have performed live trials, and using videotaped footage allowed us to collect data from a larger sample.

Subjects were provided with the following instructions at the outset of the trial:

”To help us refine human-robot interactions, we need to know exactly what people prefer or actively dislike. This trial aims to explore some important aspects of human preferences toward different robot appearances and behaviour styles. A robot companion within the home would need to know how to attract a person’s attention for different situations, and what people’s preferences are.

You will view some videotaped clips that depict a scenario where a person is busy at home, when the doorbell rings. The robot companion goes to answer the door and lets the person in, and then needs to let the person at home know that they have a visitor. The video clips will show the robot with three different appearance styles, and the ability to use different cues (e.g. lights, noises, voices) to attract your attention, in the hope of initiating an interaction with you.

We would like you to watch each video clip carefully and imagine that you are the person interacting with the robot. We would like you to tell us about your preferences by completing the questionnaire at the end of the clips.”

Once subjects had watched the three robot videotaped scenarios, they completed a robot “appearance and cues” questionnaire and a subject personality questionnaire. To date we have carried out the trials with five different groups resulting in a sample size of 80 people. The presentation of the three appearance styles video clips was counterbalanced for each of the sessions. As these trials were only completed on 16th December 2005, detailed analysis has yet to be carried out as part of the third implementation plan.

6.6 How can a robot attract people’s attention?

We assessed strategies for simple orienting responses of a human-sized robot and their impact on attracting attentional responses of humans moving in a naturalistic setting involving naive humans

moving past it in a corridor. A minimalist set-up using sonar sensors was used to recognize human movements. The robot distinguished objects from humans by assuming that only people move by themselves. Two methods using either hand-coded rules or Hidden Markov Models were implemented and compared on-line and off-line, with the robot classifying different types of movement (with respect to the robot) and interactions video-taped to provide a basis for judging if persons were interested in an interaction, depending on the type of orienting behaviour of the robot (movement orientating toward or away from the human of different magnitudes). The impact of different orienting cues by the robot in relation to the direction of from which humans approach it was studied, with regard to elicitation interaction, and results yielded insights into the importance of visibility of the robot's response to humans, whose natural activities tend not to be focused on robots, e.g. when they are merely walking past on their way elsewhere, or otherwise engaged in other activities such as interaction with other humans.

Further details about this study can be found in [7].

6.7 Future work

A number of research work will be carried out during M25-M42/48 in relation to WP6.4 including:

- Full detailed analysis of the robot appearance style studies carried out in December 2005 investigating how robot appearance and behaviour influences peoples' perceptions (Months 25-30),
- Summary of results and development of design guidelines for robot behaviour and appearance relevant to the Key Experiments (results to be disseminated to partners),
- Investigate and validate the design guidelines as part of a long-term study scheduled to commence in April 2006,
- Dissemination of results at conferences and through scientific peer-reviewed publications,
- The emphasis during 2007 (the final project year) will be on the scientific integration and validation of the results in conjunction with the KE's.

7 Future Work

We will pursue our investigation and refinement of the conceptual architecture in relation with advances and findings obtained within the project or in the state of the art. This is a continual task towards a better understanding of artificial cognition, and its links and interactions with human cognition.

In parallel, we plan to develop an interactive task planning and execution system that integrates guidelines provided by the user studies.

The next phase will involve a strong interaction between supervision and multi-modal dialog (RA1) as well as human activity interpretation (RA2) in order to effectively implement H-R interactive task achievement. Interaction with RA3 has already been implemented and will be enforced.

Development of integration and programming tools as well as relations with the other RAs will be essentially implemented in the framework of the key experiments: imitation learning (RA4) in KE3, objects and place learning and manipulation in KE1 and KE2.

8 References

9 References

9.1 Applicable documents

Relevant Publications included as appendices are:

- [1] R. Alami, A. Clodic, V. Montreuil, E. A. Sisbot, and R. Chatila. Task planning for human-robot interaction. *The Smart Objects and Ambient Intelligence Conference, sOc-EuSAI 05, Grenoble, France, 2005*.
- [2] Aurélie Clodic, Vincent Montreuil, Rachid Alami, and Raja Chatila. A decisional framework for autonomous robots interacting with humans. *IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN), 2005*.
- [3] K. Dautenhahn, S. Woods, C. Kaouri, M. Walters, K. L. Koay, and I. Werry. What is a robot companion - Friend, assistant or butler?. *Proc. IEEE IROS, Edmonton, Canada, pp. 1488-1493, 2005*.
- [4] Steffen Knoop, Sven R. Schmidt-Rohr, Rudiger Dillmann. A Flexible Task Knowledge Representation for Service Robots.
- [5] S. Woods, K. Dautenhahn, C. Kaouri, R. T. Boekhorst, and K. L. Koay. Is this robot like me? Links between human and robot personality traits. *Proc. IEEE-RAS Int. Conference on Humanoid Robots (Humanoids 2005), Tsukuba, Japan, 2005*.
- [6] M. L. Walters, K. Dautenhahn, R. Te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. Nehaniv, D. Lee, and I. Werry. The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. *Proc. Ro-Man 2005, Nashville, USA, pp. 347-352, 2005*.
- [7] M. Finke, K. L. Koay, K. Dautenhahn, C. Nehaniv, M. L. Walters, and J. Saunders. Hey, I'm over here - How can a robot attract people's attention? *Proc. IEEE Ro-Man05, Nashville, USA, pp. 7-12, 2005*.
- [8] K. L. Koay, M. L. Walters, and K. Dautenhahn. Methodological issues using a comfort level device in human-robot interactions. *Proc. IEEE Ro-Man05, Nashville, USA, pp. 359-364, 2005*.
- [9] K. L. Koay, K. Dautenhahn, S. Woods, and M. L. Walters. Empirical results from using a comfort level device in human-robot interaction studies. *To appear in Proc. HRI06, Salt Lake City, Utah, USA, 2006*.
- [10] J. Fritsch, M. Kleinhagenbrock, A. Haasch, S. Wrede, and G. Sagerer. A flexible infrastructure for the development of a robot companion with extensible hri-capabilities. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2005*.
- [11] Shuyin Li, Axel Haasch, Britta Wrede, Jannik Fritsch, and Gerhard Sagerer. Human-style interaction with a robot for cooperative learning of scene objects. In *Proc. Int. Conf. on Multimodal Interfaces*, pages 151–158, Trento, Italy, 2005. ACM Press.

Further submitted publications:

- [12] S. N. Woods, M. L. Walters, K. L. Koay, K. Dautenhahn. Comparing Human Robot Interaction Scenarios Using Live and Video Based Methods: Towards a Novel Methodological Approach, *accepted for AMC06 (Advanced Motion Control), Istanbul from March 27-29 2006*, submitted and accepted.
- [13] S. Woods, K. Dautenhahn, C. Kaouri, R. T. Boekhorst, K. L. Koay, and M. Walters. Are robots like people? The role of subject and robot personality traits in human-robot interaction trials. *International Journal of Humanoid Robotics*, submitted.
- [14] R. Alami, A. Clodic, V. Montreuil, E. A. Sisbot, and R. Chatila. Toward Human-Aware Robot Task Planning. *AAAI 2006 Spring Symposium "To Boldly Go Where No Human-Robot Team Has Gone Before" 06, Stanford, USA*, submitted and accepted.

9.2 Reference documents

- [15] Rachid Alami, Raja Chatila, Sara Fleury, Malik Ghallab, and Flix Ingrand. An architecture for autonomy. *International Journal of robotics Research, Special Issue on Integrated Architectures for Robot Control and Programming*, 17(4), 1998.
- [16] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglas, C. Lebiere, and Y. Qin. An integrated theory of the mind. *Psychological Review*, 111(4):1036–1060, 2004.
- [17] R. A. Arkin, and T. Balch. AuRA: Principles and Practice in Review. *Journal of Experimental and Theoretical AI*, 9(2), 1997.
- [18] Winfried Baum, Christopher Parlitz. Go: Description and Manual. *unpublished*.
- [19] R. Peter Bonasso, D. Kortenkamp, D. Miller, and M. Slack. Experiences with an architecture for intelligent reactive agents. *Journal of Experimental and Theoretical AI*, 9(2), 1997.
- [20] Cynthia Breazeal (Ferrell). A motivational system for regulating human-robot interaction. *National Conference of American Association for Artificial Intelligence (AAAI)*, 1998.
- [21] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2(1):14–23, March 1986.
- [22] Daphna Buchsbaum, Bruce Blumberg, Cynthia Breazeal, and Andrew N. Meltzoff. A simulation-theory inspired social learning system for interactive characters. *IEEE International Workshop on Robots and Human Interactive Communication*, 2005.
- [23] Aurélie Clodic, Sara Fleury, Rachid Alami, Matthieu Herrb, and Raja Chatila. Supervision and interaction: Analysis from an autonomous tour-guide robot deployment. *In the 12th International Conference on Advanced Robotics*, 2005.
- [24] Philip R. Cohen and Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–361, 1990.
- [25] Philip R. Cohen and Hector J. Levesque. Teamwork. *Nous*, 25(4):487–512, 1991.

- [26] David Feil-Seifer and Maja J. Mataric. A multi-modal approach to selective interaction in assistive domains. *IEEE International Workshop on Robots and Human Interactive Communication*, pages 416–421, 2005.
- [27] S. Fleury, M. Herrb, and R. Chatila. Genom: A tool for the specification and the implementation of operating modules in a distributed robot architecture. In *International Conference on Intelligent Robots and Systems*, volume 2, pages 842–848, Grenoble (France), 1997.
- [28] Terrence W. Fong, Illah Nourbaksh, Clayton Kunz, Lorenzo Flckiger, John Schreiner, Robert Ambrose, Robert Burrige, Reid Simmons, Laura M. Hiatt, Alan Schultz, J. Gregory Trafton, Magda Bugajska, and Jean Scholtz. The peer-to-peer human-robot interaction project. *AIAA Space 2005*, September 2005.
- [29] Erann Gat. Integrating reaction and planning in a heterogeneous asynchronous architecture for mobile robot navigation. *SIGART Bulletin*, 2(4):70–74, 1991.
- [30] F. Gravot, S. Cambon, and R. Alami. asymov: a planner that deals with intricate symbolic and geometric problem. *Springer Tracts in Advanced Robotics, Volume 15 / 2005, Robotics Research: The Eleventh International Symposium*, 2005.
- [31] Ian Horswill. Tagged behavior-based systems: Integrating cognition with embodied activity. *IEEE Intelligent Systems, September/October 2001*, pp 30-38, 2001.
- [32] Kazuhiko Kawamura, Phongchai Nilas, and Kazuhiko Mugumura. An agent-based architecture for an adaptive human-robot interface. *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS)*, 2003.
- [33] Sanjeev Kumar, Marcus J. Huber, and Philip R. Cohen. Representing and executing protocols as joint actions. *Proceedings of the First International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS)*, 2002.
- [34] J. F. Lehman, J. Laird, and P. Rosenbloom. A gentle introduction to soar : An architecture for human cognition: 2006 update. <http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf>, 2006.
- [35] Hector J. Levesque, Philip R. Cohen, and J. H. T. Nunes. On acting together. *Proceedings of the Annual Meeting of the American Association for Artificial Intelligence (AAAI)*, 1990.
- [36] M. Likhachev, M. Kaess, and R. Arkin. Learning behavioral parametrization using spatio-temporal case-based reasoning. In *IEEE International Conference on Robotics and Automation*, 2002.
- [37] R. T. Maheswaran, M. P. Tambe, Varakantham, and K. Myers. Adjustable autonomy challenges in personal assistant agents: A position paper. *Proceedings of Autonomy'03*, 2003.
- [38] Hiroshi Nakajima, Yasunori Morishima, Scott Brave, Ryota Yamada, Heidi Maldonado, Clifford Nass, Masaki Arai, and Shigeyasu Kawaji. Toward an actualization of social intelligence in human and robot collaborative systems. *International Conference on Intelligent Robots and Systems (IROS)*, 2004.

- [39] Dana Nau, Tsz-Chiu Au, Okhtay Ilghami, Ugur Kuter, J. William Murdock, Dan Wu, and Fusun Yaman. Shop2: An htn planning system. *Journal of Artificial Intelligence Research*, (20):379–404, december 2003.
- [40] A. Newell. *Unified Theories of Cognition*. Harvard University Press, Cambridge, Massachusetts, 1990.
- [41] S. Nonaka, K. Inoue, T. Arai, and Y. Mae. Evaluation of human sense of security for coexisting robots using virtual reality. *IEEE International Conference on Robotics and Automation (ICRA)*, 2004.
- [42] Illah Nourbakhsh, Katia Sycara, Mary Koes, Mark Yong, Michael Lewis, and Steve Burion. Human-robot teaming for search and rescue. *IEEE Pervasive Computing: Mobile and Ubiquitous Systems*, pages 72–78, 2005.
- [43] Illah R. Nourbakhsh and Terrence Fong. Human-robot teams on the moon: Challenges and plans (extended abstract). *International Conference on Robotics and Automation (ICRA), Workshop on Fielding Multi-Robot Systems*, 2005.
- [44] F. Py and F. Ingrand. Dependable execution control for autonomous robots. In *IROS 2004 (IEEE/RSJ International Conference on Intelligent Robots and Systems)*, Sendai, Japan, 2004.
- [45] Python Programming Language: <http://www.python.org>
- [46] Rosenbloom P., Laird J., and Newell A. eds. *The Soar Papers: Research on Integrated Intelligence*. MIT Press, Cambridge, Massachusetts, 1993.
- [47] Simple API for XML (SAX): <http://www.saxproject.org>
- [48] Paul Scerri, David Pynadath, Nathan Schurr, Alessandro Farinelli, Sudeep Gandhe, and Milind Tambe. Team oriented programming and proxy agents: The next generation. In *Proceedings of 1st international workshop on Programming Multiagent Systems*, 2004.
- [49] Paul Scerri, David V. Pynadath, Lewis Johnson, Rosenbloom Paul, Nathan Schurr, Mei Si, and Milind Tambe. A prototype infrastructure for distributed robot-agent-person teams. *The Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2003.
- [50] Nathan Schurr, Paul Scerri, and Millind Tambe. Coordination advice: A preliminary investigation of human advice to multiagent teams. *AAAI Spring Symposium on Interaction between Humans and Autonomous Systems over Extended Operation, (Invited Paper)*, 2004.
- [51] Reid Simmons. Structured control for autonomous robots. *IEEE Transactions on Robotics and Automation*, 10(1), February 1994.
- [52] Milind Tambe. Agent architectures for flexible, practical teamwork. *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 1997.
- [53] J. Gregory Trafton, Nicholas L. Cassimatis, Magdalena D. Bugajska, Derek P. Prock, Farilee E. Mintz, and Alan C. Schultz. Enabling effective human-robot interaction using perspective-taking in robots. *IEEE Transactions on Systems, Man and Cybernetics*, 35(4):460–470, 2005.

- [54] J. Gregory Trafton, Alan C. Schultz, Magdalena Bugajska, and Farilee Mintz. Perspective-taking with robots: Experiments and models. *IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN)*, 2005.
- [55] Ulrich Greeve. Entwicklung zweier Komponenten einer globalen Ablaufsteuerung für einen mobilen Haushaltsroboter *Diploma Thesis Fraunhofer IPA*, 2005.
- [56] R. Volpe, I.A.D. Nesnas, T. Estlin, D. Mutz, R. Petras, and H. Das. The clarity architecture for robotic autonomy. In *IEEE 2001 Aerospace Conference*, March 2001.

10 Annex: papers related to RA6

1. R. Alami, A. Clodic, V. Montreuil, E. A. Sisbot, and R. Chatila. Task planning for human-robot interaction. *The Smart Objects and Ambient Intelligence Conference, sOc-EuSAI 05, Grenoble, France, 2005.*
2. Aurélie Clodic, Vincent Montreuil, Rachid Alami, and Raja Chatila. A decisional framework for autonomous robots interacting with humans. *IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN), 2005.*
3. K. Dautenhahn, S. Woods, C. Kaouri, M. Walters, K. L. Koay, and I. Werry. What is a robot companion - Friend, assistant or butler?. *Proc. IEEE IROS, Edmonton, Canada, pp. 1488-1493, 2005.*
4. M. Finke, K. L. Koay, K. Dautenhahn, C. Nehaniv, M. L. Walters, and J. Saunders. Hey, I'm over here - How can a robot attract people's attention? *Proc. IEEE Ro-Man05, Nashville, USA, pp. 7-12, 2005.*
5. J. Fritsch, M. Kleinhagenbrock, A. Haasch, S. Wrede, and G. Sagerer. A flexible infrastructure for the development of a robot companion with extensible hri-capabilities. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2005.*
6. K. L. Koay, M. L. Walters, and K. Dautenhahn. Methodological issues using a comfort level device in human-robot interactions. *Proc. IEEE Ro-Man05, Nashville, USA, pp. 359-364, 2005.*
7. K. L. Koay, K. Dautenhahn, S. Woods, and M. L. Walters. Empirical results from using a comfort level device in human-robot interaction studies. *To appear in Proc. HRI06, Salt Lake City, Utah, USA, 2006.*
8. Steffen Knoop, Sven R. Schmidt-Rohr, Rudiger Dillmann. A Flexible Task Knowledge Representation for Service Robots.
9. Shuyin Li, Axel Haasch, Britta Wrede, Jannik Fritsch, and Gerhard Sagerer. Human-style interaction with a robot for cooperative learning of scene objects. In *Proc. Int. Conf. on Multimodal Interfaces*, pages 151–158, Trento, Italy, 2005. ACM Press.
10. M. L. Walters, K. Dautenhahn, R. Te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. Nehaniv, D. Lee, and I. Werry. The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. *Proc. Ro-Man 2005, Nashville, USA, pp. 347-352, 2005.*
11. S. Woods, K. Dautenhahn, C. Kaouri, R. T. Boekhorst, and K. L. Koay. Is this robot like me? Links between human and robot personality traits. *Proc. IEEE-RAS Int. Conference on Humanoid Robots (Humanoids 2005), Tsukuba, Japan, 2005.*

Task planning for Human-Robot Interaction

Rachid Alami, Aurélie Clodic, Vincent Montreuil, Emrah Akin Sisbot, Raja Chatila

LAAS - CNRS

7, Avenue du Colonel Roche 31077 Toulouse, France

Firtsname.Name@laas.fr

Abstract— Human-robot interaction requires explicit reasoning on the human environment and on the robot capacities to achieve its tasks in a collaborative way with a human partner.

This paper focuses on organization of the robot decisional abilities and more particularly on the management of human interaction as an integral part of the robot control architecture. Such an architecture should be the framework that will allow the robot to accomplish its tasks but also produce behaviors that support its engagement vis-a-vis its human partner and interpret similar behaviors from him.

Together and in coherence with this framework, we intend to develop and experiment various task planners and interaction schemes, that will allow the robot to select and perform its tasks while taking into account explicitly the constraints imposed by the presence of humans, their needs and preferences.

We have considered a scheme where the robot plans for itself and for the human in order not only (1) to assess the feasibility of the task (at a certain level) before performing it, but also (2) to share the load between the robot and the human and (3) to explain/illustrate a possible course of action.

I. INTRODUCTION

The introduction of robots in our daily life raises a key issue that is “added” to the “standard challenge” of autonomous robots: the presence of humans in its environment and the necessity to interact with them. Clearly, the human should be taken explicitly into account in all steps of the robot design.

We are conducting research on robot decisional abilities taking into account explicit reasoning on the human environment and on the robot capacities to achieve its tasks in such a context.

This paper focuses on organization of the robot decisional abilities and more particularly on the management of human interaction as an integral part of the robot control architecture. Such an architecture should be the framework that will allow the robot to accomplish its tasks but also produce behaviors that support its engagement vis-a-vis its human partner and interpret similar behaviors from him.

Together and in coherence with this framework, we intend to develop and experiment various task planners and interaction schemes that will allow the robot to select and perform its tasks while taking into account explicitly the constraints imposed by the presence of humans, their needs and preferences.

Section II discusses briefly related work. Section III draws a general view of the framework that we propose. In section IV and V we discuss new human-interaction related issues in symbolic action planning as well as in

motion planning. The last section presents an application that will serve as an implementation and validation testbed.

II. RELATED WORK

A number of contributions in Human-Robot Interaction (HRI) involve a human operator who controls the robot from a distant place [17], [15], [1]. Besides tele-operation issues, the main issues that are treated in this context are mixed initiative, shared decision and adjustable autonomy. Indeed, in such context the human intervenes essentially at interpretation and decision level.

In our context the human is physically present in the vicinity of the robot, is sensed by the robot and may even participate to the task performance. In such applications, HRI takes place at different levels [14]: verbal, visual, physical, decisional, etc. . .

Only a limited number of papers consider the robot and the human as agents who can cooperate to achieve common goals. The current paper focuses on this particular issue. One major point is that the robot must act in a way judged as “acceptable” by humans.

In relation with this, a number of recent contributions about close interaction deal with the notion of physical and mental safety [22] or the introduction of emotions and/or cognitive models in robotic structures [6], [21].

Very often, HRI is merged into the task performance. This tends to reduce HRI to a (sometimes very sophisticated) human interface.

Our aim is to endow the robot with an explicit consideration of humans and with the ability to manage its interactions with them. This must be considered at different levels: at the architecture level as well as at the task planning and motion planning levels.

Our first source of inspiration is the Joint Intention theory (see [11], [19], [12]). It is based on the notion of commitment for team members and defines for a team the concept of Joint Persistent Goal. These definitions constitute a basis for the elaboration of cooperation schemes between heterogeneous agents (see [16] for an example). However these definitions are very general and we have tried to adapt them to our context.

One problem in the design of an architecture for HRI is the representation of humans. In fact, the attitude of a human depends on a great number of factors more or less controllable. A good idea is the representation of an agent with a proxy. This approach has been explored and implemented in STEAM and more recently in Machinetta (see [28], [24], [25]).

Our robot is controlled by a three layer architecture (see [2]). We discuss here below the design of the decisional level in which we introduce what we call InterAction Agents (IAAs). They are similar to proxies but are directly implemented on the robot side as a representative of a human agent. To make the interaction more explicit we have defined a complete process of establishing a common goal, achieving it and verifying commitment of all agents involved. Besides, relevant IAA models should be devised and used in the robot planning activities. Such models will range from high-level specifications of the human abilities and preferences to geometric attributes such as position, posture or visibility regions.

III. DECISIONAL SYSTEM FRAMEWORK

We envision HRI in a context where two agents (a human and a robot) share a common space and exchange information through various modalities[10].

Interaction happens as a consequence of an explicit request of the human to satisfy a goal or because the robot finds itself in a situation where it is useful if not mandatory.

In both cases, the robot has a goal to satisfy. An important issue is the notion of engagement, a process in which the robot will have to establish, maintain and end a connection with a human partner. Besides conversation, such a process will provide a framework for robots performing tasks in a human context.

This will cover goal establishment, selection of an incremental refinement of the task that is intended to achieve it. The establishment of a connection between the human and the robot will serve to the robot follow human task performance and to monitor his/her commitment to the common goal, and even to influence it.

In order to deal with the various aspects that the decisional kernel of the robot has to do, we have designed a decisional framework which consists of several entities, having each a specific role. The global view is illustrated by Fig. 1.

The HRI we consider in this paper is the common achievement of tasks by two agents - a robot and a human - in order to satisfy a joint goal.

The Agenda: Several goals may be sought at a given time, involving possibly several persons. At any moment, there may be several active, inactive and suspended goals. The Agenda manages the current set of robot goals. It ensures the consistency between active goals, and determines their priorities, and their causal links. Based on data provided by the Supervision Kernel, the Agenda determines the relevance of goals and decides to create, suspend, resume or abandon a goal. When a goal is created, it may be associated to the robot alone or to a “team” of agents.

The IAA Manager: The humans encountered by the robot are represented by entities called “InterAction Agents” (IAAs). An IAA is created dynamically and maintained by the “IAA Manager”.

The Task Delegates: The set of active goals entails the incremental execution of a set of tasks in interaction with humans. Each task corresponding to an active or a

suspended goal is represented by an entity called “Task Delegate” that is in charge of monitoring the progress towards the goals of both the robot and the IAA and to assess the level of commitment of the associated person. To do so, it controls a set of “Observers” (OBs).

The Robot Supervision Kernel: The Robot Supervision Kernel is responsible of all tasks selection, refinement and execution. It maintains an integrated view of all robot activities and ensures a global coherence of robot behavior. It is the only entity that can send execution requests to the functional level.

For each new active goal the Robot Supervision Kernel creates a Task Delegate, selects or elaborates a plan and allocates the roles of each team member. Notice that the creation of the Task Delegate is combined with the creation of an OB for each human involved in the task performance.

For all the other active goals, the Robot Supervision Kernel has already a plan and is in charge of the execution of the robot part. Whenever an elementary action is performed, the Robot Supervision Kernel forwards this information to all active Tasks Delegates.

Depending on the context, the planning process can be more or less elaborated. Indeed, the presence of humans in the environment raises new issues in the classic motion, manipulation and task planning. We are developing, in coherence with the architecture presented here, a motion planner [27], [23] that can be used not only to plan safe robot paths, but also to plan good, socially acceptable and legible paths and a high-level task planner [20], [10] that is able to deal with constraints imposed by the presence of humans, their needs and preferences.

IV. HIGH-LEVEL SYMBOLIC PLANNING

Context: The main point here is how high level robot planning skills should be developed in order to allow it to act as a companion.

In such a scheme, the robot plans for itself and for the human in order:

- not only, to assess the feasibility of the task (at a certain level) before performing it
- but also, to share the load between itself and the human
- and also, to explain/illustrate a possible course of actions.

We concentrated on a planner that is able to take into account “social constraints” and to synthesize plans compatible with human preferences, acceptable by humans and easily legible in terms of intention.

Representing social constraints: We have elaborated a formalization where both the robot and the human are represented in terms of actions they can perform. In a first tentative, we have limited our representation to STRIPS-like domains.

A “team” composed of two “actors” (the robot and a human) can be represented as: $(A_{human}, C_{human}^{ctxt})$ and $(A_{robot}, C_{robot}^{ctxt})$ where A_i are sets of actions and C_i^{ctxt} are their context-dependent associated costs. The costs

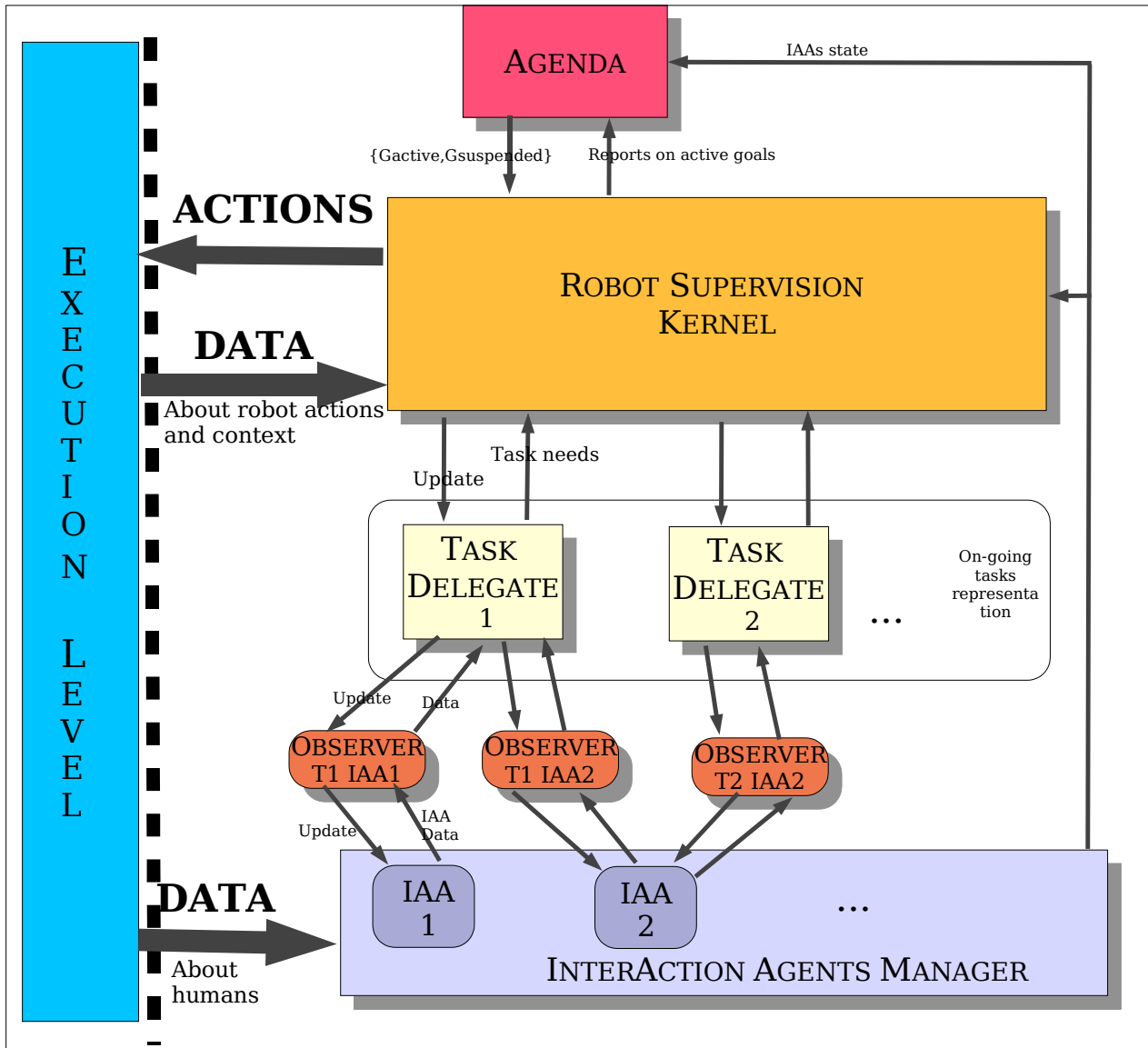


Fig. 1. Decisional framework for a HRI-enabled robot

represent the difficulty and the pleasure an actor has in an action realization.

Besides, in order to take into account issues linked to the acceptability of a plan by a human, we associate a cost to certain situations and to certain actions sequences in order to model states and action courses that might be unacceptable or inconvenient for the human.

Preliminary tests have been conducted based on a HTN (Hierarchical Task Networks) planner SHOP2[1] mainly because it permits to specify costs for each action and it can produce plans with the least total cost.

Examples involved domestic like situations where the robot essentially performs fetch-and-carry and cleaning tasks in interaction with a human. This study have confirmed [20], for us, the relevance of this level and of the types of considerations that should be taken into account

when building robot plans in this context. This should be the basis for task planning but also, as we mentioned, role allocation, dialogue about plans, human-robot 'negotiation'.

V. HUMAN-AWARE MOTION PLANNING

The presence of humans in the environment raises also new issues to the classic motion-manipulation-task planning [8], [23].

We claim that a human-aware motion planner must not only elaborate safe robot paths, but also plan good, socially acceptable and legible paths. Our aim is to build a planner that takes explicitly into account the human partner by reasoning about his accessibility, his vision field and potential shared motions.

While several contributions take into account safety criteria (distance, inertia), very few papers, in our knowledge,

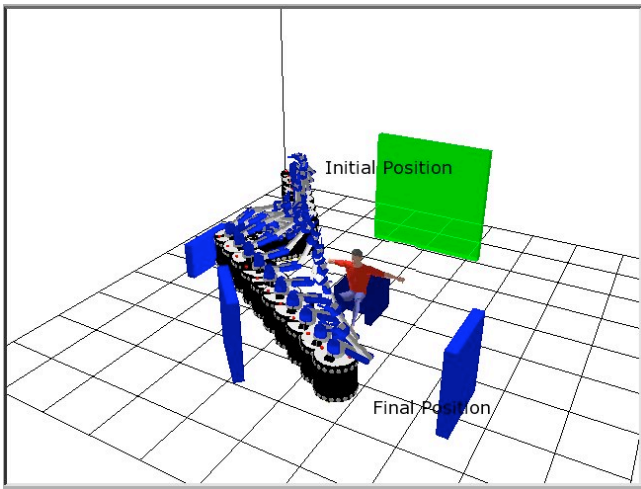


Fig. 2. A path produced by a conventional planner

deal with comfort and legibility issues and often in an ad hoc manner. We believe that our approach can be more generic. We introduce two criteria to the motion planning stage to ensure safety and comfort. The robot must take into account these two criteria at the planning stage along with the more common aspects of path planning, i.e. obstacle avoidance and shortest path finding. The first criterion, called security criterion, mainly focuses on ensuring the safety by controlling the distance between robot and human. The robot, if possible, must avoid approaching too much to humans, and in some cases a certain perimeter around humans must not be allowed to pass through. The sudden appearance of the robot from behind an obstacle may cause fear and surprise especially if the obstacle is close to the human.

Another criterion, called visibility criterion, takes into account the human's field of view and robot's relative position relatively to it. Humans tend to feel safer and more comfortable when the robot is in their sight. It is preferable that the robot chooses a path as visible as possible to ensure this criterion. The visible and invisible zones can be ranked proportionally to the minimum angular deviation from the human's gaze. Indeed, we can consider this visibility criterion that is proportional to the "human's effort to keep the robot in his sight by turning the head or the body".

Note that other aspects should be taken into account like speed (time to contact) and acceleration of the robot (or a part of it) particularly when it is in the close vicinity of the persons.

We are investigating various minimization criteria based on a weighted combination of distance, visibility and comfort for computing a satisfactory path and velocity profile. Preliminary results with a comparison to the conventional planner are shown in Fig. 3.

VI. AN APPLICATION CONTEXT

One application on which we envisage to implement and test the proposed approach is an interactive tour-guide robot

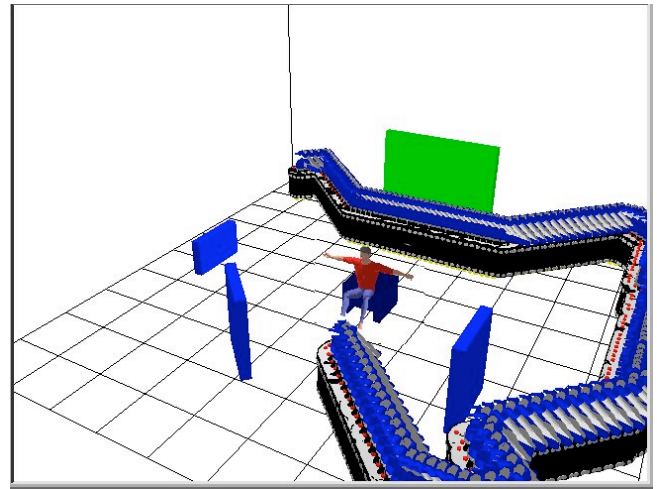


Fig. 3. A path that takes explicitly into account the human presence and preferences. The algorithm conducts the robot to approach the human in face to face situation.

called Rackham (Fig. 4). Let us first briefly introduce it.

Rackham: Rackham[9] has been designed as a new tour-guide robot. Besides robustness and efficiency in the robot basic navigation abilities in a dynamic environment, our focus was to develop and test a methodology to integrate HRI abilities in a systematic way.

To test and validate our developments, we have decided to bring it regularly (two weeks every three months) to a museum in Toulouse. Rackham has already been used at the exhibition for hundreds of hours (May 2004, July 2004, February 2005, May 2005), accumulating valuable data and information for future enhancements. The project is conducted so as to incrementally enhance the robot functional and decisional capabilities based on the observation of the interaction between the public and the robot.

A number of features have been installed for HRI:

- the detection of dynamic "obstacles",
- a vision-based face detector[7],
- a 3D animated head with speech synthesis[4],
- displays and inputs from the touch screen,
- control of robots lights.

In its current version, the emphasis has been mainly put on robustness in a dynamic environment. All HRI features currently running on Rackham have been classically encoded as event-driven automata with no explicit management of the interactions and no reasoning on human behavior. The next step is to implement the proposed framework for HRI.

Rackham desired capabilities: Here are some examples of the desired abilities:

- when left alone, Rackham should seek for people to interact with.
- Rackham should be able to detect various types of persons and adapt its behavior to them.
- Rackham should be able to manage two or more interactions in parallel.
- Rackham should be able to measure level of com-

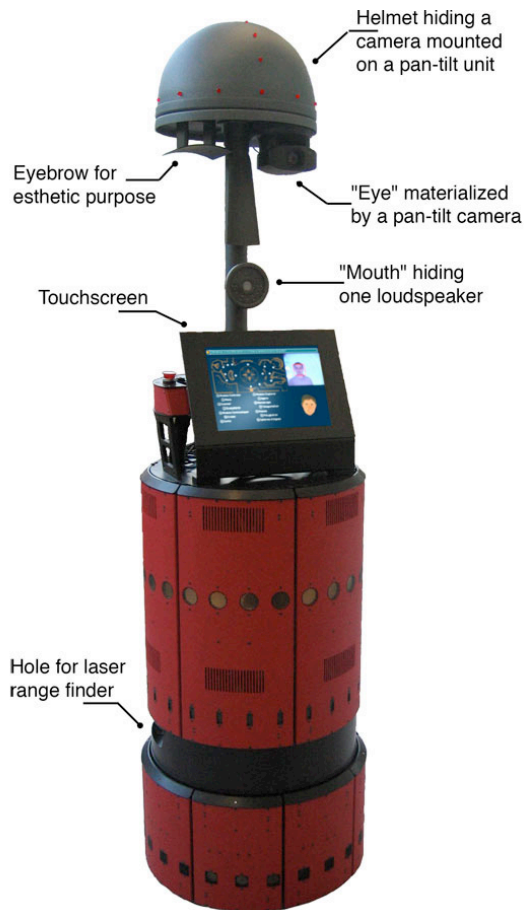


Fig. 4. Rackham: a tour guide robot

mitment of its human interactors and should react accordingly; for instance, detecting that the guided person follows slowly or is no more interested by the tour.

There will be various types of IAA corresponding to the different types of persons that Rackham might encounter: *passer-by*, *visitor*, *operator* will have their specific abilities and preferences. Rackham will behave and interact differently with them.

VII. CONCLUSION AND FUTURE WORK

In this paper we have presented a decisional framework designed for robots operating in a human environment. Our objective was to provide a management of human interaction that can be seen as an integral part of a general robot control architecture. This was done in order to provide a principled way to deal with HRI.

We also intend to use the developed approach as a framework in which we will develop and experiment various task planners and interaction schemes that explicitly consider human abilities and preferences.

The next steps will be a further refinement of the framework proposed here and its implementation on a physical robot.

ACKNOWLEDGMENTS

The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020 and by the French National Program ROBEA.

Numerous persons have contributed to Rackham: the authors want to thanks: Aymeric Yvernes, Coline Lelong-Pantel, Stéphane Plais, Guillaume Infantes, Maxime Cottret, Paulo Menezes, Ludovic Brethes, Frederic Lerasle; from the Space City Museum : Jean-Noël Plachez, Véronique Hallard, Maxime Randisi and Stephan Lauriol; and from ICP Grenoble : Gérard Bailly and Frédéric Elisei.

REFERENCES

- [1] J. Allen, G. Ferguson. "Human-Machine Collaborative Planning", *In Proceedings of the Third International NASA Workshop on Planning and Scheduling for Space*, 2002.
- [2] R. Alami, R. Chatila, S. Fleury, M. Ghallab, F. Ingrand, "An architecture for autonomy", *International Journal of Robotic Research*, Vol.17, N°4, pp.315-337, April 1998.
- [3] R. Alami, T. Simeon, and K. Madhava Krishna. "On the influence of sensor capacities and environment dynamics onto collision-free motion plans." *Proc. in IEEE International Conference on Intelligent Robots and Systems*, 2002.
- [4] G. Bailly, M. Berar, F. Elisei, and M. Odisio, "Audiovisual speech synthesis", *International Journal of Speech Technology*, 6:331-346, 2003.
- [5] C. Breazeal, G. Hoffman, and A. Lockerd, "Teaching and Working with Robots as a Collaboration," *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2004.
- [6] C. Breazeal, "A Motivational System for Regulating Human-Robot Interaction," *Proceedings of National Conference of American Association for Artificial Intelligence (AAAI)*, 1998.
- [7] L. Brethees, P. Menezes, F. Lerasle, and J. Hayet, "Face tracking and hand gesture recognition for human robot interaction," *International Conference on Robotics and Automation (ICRA)*, New Orleans, pp. 1901-1906, 2004.
- [8] R. Chatila, R. Alami, T. Simeon, J. Pettre, L. Jaillet, "Safe, reliable and friendly interaction between humans and humanoids", *Rapport LAAS No02530, 3rd IARP International Workshop on Humanoid and Human Friendly Robotics*, Tsukuba (Japan), 11-12 December 2002, pp.83-87
- [9] A. Clodic, S. Fleury, R. Alami, M. Herrb, R. Chatila, "Supervision and Interaction: Analysis from an Autonomous Tour-guide Robot Deployment," *International Conference on Advanced Robotic (ICAR)*, 2005.
- [10] A. Clodic, V. Montreuil, R. Alami, R. Chatila, "A Decisional Framework for Autonomous Robots Interacting with Humans," *14th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 05)*, Nashville, USA, August 2005.
- [11] P. Cohen and H. Levesque, "Intention is Choice with Commitment," *Artificial Intelligence*, vol. 42, pp. 213-261, 1990.
- [12] P. R. Cohen, and H. J. Levesque. "Teamwork". *Nous*, Special Issue on Cognitive Science and AI, 25, 25(4): 487-512, 1991.
- [13] G. A. Fink, J. Fritsch, S. Hohener, M. Kleinhagenbrock, S. Lang, and G. Sagerer. "Towards multi-modal interaction with a mobile robot". *Pattern Recognition and Image Analysis*, 14(2):173-184, 2004.
- [14] T. Fong, I. Nourbakhsh, K. Dautenhahn. "A survey of socially interactive robots," *Robotics and Autonomous Systems*, *Special issue on Socially Interactive Robots*, 42 (3-4), pp 143-166, 2003.
- [15] T.W. Fong, C. Thorpe, and C. Baur, "Collaboration, Dialogue, and Human-Robot Interaction," *Proceedings of the 10th International Symposium of Robotics Research*, Lorne, Victoria, Australia, Springer-Verlag, London, November, 2001.
- [16] N. Jennings, "Specification and Implementation of a Belief- Desire-Joint-Intention Architecture for Collaborative Problem Solving," *International Journal of Intelligent and Cooperative Information Systems*, vol. 2, no. 3, pp. 289-318, 1993.

- [17] K. Kawamura and P. Nilas and K. Muguruma and J. A. Adams and C. Zhou, "An Agent-Based Architecture for an Adaptive Human-Robot Interface", *Hawaii International Conference on System Sciences (HICSS'03)*, 2003.
- [18] D. Kulic and E. Croft. "Safe planning for human-robot interaction", *Proc. in IEEE International Conference on Robotics & Automation, New Orleans, LA*, 2004.
- [19] H. J. Levesque, P. R. Cohen, and J. H. T. Nunes. "On acting together." *In Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, pages 94-99, 1990.
- [20] V. Montreuil, R. Alami "Report on Paradigms for Decisional Interaction" *Cogniron Deliverable D6.2.1, LAAS Internal Report*, January 2005.
- [21] H. Nakajima, et al, "Toward an Actualization of Social Intelligence in Human and Robot Collaborative Systems," *International Conference on Intelligent Robots and Systems (IROS)*, 2004.
- [1] D. Nau, T.-C. Au, O. Ilghami, U. Kuter, W. Murdock, D. Wu, and F.Yaman. *SHOP2: An HTN Planning System*. JAIR, volume 20, pp. 379-404, 2003.
- [22] S. Nonaka, K. Inoue, T. Arai and Y. Mae, "Evaluation of Human Sense of Security for Coexisting Robots using Virtual Reality," *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2004.
- [23] E. A. Sisbot, R. Alami and T. Simeon, K. Dautenhahn, M. Walters, S. Woods, K. L. Koay and C. Nehaniv, " Navigation in the Presence of Human", IEEE-RAS International Conference on Humanoid Robots, Humanoids2005, Tsukuba, Japan, December 2005.
- [24] P. Scerri, et al, "A Prototype Infrastructure for Distributed Robot-Agent-Person Teams," *The Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2003.
- [25] P. Scerri, et al, "Team Oriented Programming and Proxy Agents: The Next Generation," *Proceedings of the First International Workshop on Programming Multiagent Systems*, (2004).
- [26] Sidner, C.; Lee, C.; Kidd, C.; Lesh, N., "Explorations in Engagement for Humans and Robots", *IEEE RAS/RSJ International Conference on Humanoid Robots*, November 2004.
- [27] A. Sisbot, R. Alami "Architecture of the placement planner for human-robot close interaction" *Cogniron Deliverable D3.5.1, LAAS Internal Report*, January 2005.
- [28] M. Tambe, "Agent Architectures for Flexible, Practical Teamwork," *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 1997.

A Decisional Framework for Autonomous Robots Interacting with Humans*

Aurélie Clodic, Vincent Montreuil, Rachid Alami, Raja Chatila

LAAS - CNRS

7, Avenue du Colonel Roche 31077 Toulouse, France

{Aurelie.Clodic,Vincent.Montreuil,Rachid.Alami,Raja.Chatila}@laas.fr

Abstract—The presence of humans in its environment and the necessity to interact with them raise new challenges to the robot. Indeed, they require explicit reasoning on the human environment and on the robot capacities to achieve its tasks in a collaborative way with a human partner.

This paper focuses on architectural aspects and more particularly on the organization of the robot decisional abilities for interacting with people. Indeed, our objective is to develop a management of human interaction that will be an integral part of a general robot control architecture. This should hopefully allow to come up with a principled way to deal with human-robot interaction for robot task achievement in presence of humans or in synergy with humans.

Such an architecture should be the framework that will allow the robot to produce behaviors to accomplish its tasks but also produce behaviors that support its engagement vis-a-vis its human partner and interpret similar behaviors from him.

We also intend to use the proposed approach as a framework in which we will develop and experiment various task planners and interaction schemes. Indeed, the robot should be able, for instance, to devise plans that allow it to execute its actions and to place itself to be seen by or to observe humans, according to the task.

I. INTRODUCTION

The introduction of robots in our daily life raises a key issue that is “added” to the “standard challenge” of an autonomous robot: the presence of humans in its environment and the necessity to interact with them. Clearly, the human should be taken explicitly into account in all steps of the robot design.

We are conducting research on robot decisional abilities taking into account explicit reasoning on the human environment and on the robot capacities to achieve a task. This paper focuses on architectural aspects. Indeed, our objective is to develop a management of human interaction that will be an integral part of a general robot control architecture. This should allow to come up with a principled way to deal with Human-Robot Interaction (HRI) for task achievement in presence of humans or in synergy with humans.

Such an architecture should be the framework that will allow the robot (1) to produce behaviors to accomplish its

tasks, (2) to produce behaviors that support its engagement vis-a-vis its human partner, and (3) to interpret the human behavior relatively to the task and to itself.

We also intend to use the proposed approach as a framework in which we will develop and experiment various task planners and interaction schemes. Indeed, the robot should be able, for instance, to devise plans that allow it to execute its actions and to place itself to be seen by or to observe humans and their activity, according to the task.

Section II discusses related work briefly. Section III draws a general view of the framework that we propose. In section IV we provide an illustrated description.

II. RELATED WORK

A number of papers on HRI involve the presence of a human operator who controls the robot from a distant place [15], [13], [1]. Besides tele-operation issues, the main aspects that are treated in this context are mixed initiative, shared decision and adjustable autonomy. Indeed, in such contexts the human intervenes essentially at interpretation and decision level.

Our context assumes that the human is physically present in the vicinity of the robot, is sensed by the robot and may even participate to the task achievement. This is a very large domain [12] because HRI takes place at different levels: verbal, visual, physical, decisional, etc... All these aspects are intimately linked and it is difficult to draw boundaries between them and to treat them separately.

Only a limited number of papers consider the robot and the human as agents who can cooperate to achieve common goals. The current paper focuses on this particular issue. One major key point is that the robot must act in a way judged as *legible* and *acceptable* by humans.

In relation with this, a number of recent contributions about close interaction study the notion of physical and mental safety [20] or the introduction of emotions and/or cognitive models in robotic structures [6], [19].

Most contributions on HRI deal with robots dedicated to specific tasks. These robots are controlled by task-dependent software architectures which do not explicit clearly the notion of interaction with humans. This tends to reduce HRI to a sometimes very sophisticated human interface, which can be

*The work described in this paper was partially conducted within the EU Integrated Project COGNIRON (“The Cognitive Companion”) and funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

verbal, to select and launch pre-programmed behaviors when a human issues a request or an order. This approach is not satisfying for a robot acting in an open human environment because humans are reduced to a "clicker" and/or to an obstacle when the robot moves. Moreover, in such cases, the robot control architecture is directly derived from the application and so necessitates an important adaptation work from one application to another. We contend it is necessary to endow the robot with a more explicit consideration of humans in its environment and especially with an explicit management of interactions. This must be considered at different levels: at the planning level (social behavior, human safety, etc...), at motion planning level and at the software architecture level. In this paper, our main focus is the control architecture of robots for HRI.

In a three-layered architecture [2] including a decisional higher level, an execution control level and a functional level, we propose that the decisional level be augmented by an interaction capacity. To give a concrete expression of the humans consideration by the robot we use what we call "InterAction Agents" (IAAs) which are implemented on the robot side as representing the humans.

Our first source of inspiration is the Joint Intention theory [9], [17], [10]. It is based on the notion of commitment for team members and defines for a team the concept of "Joint Persistent Goal". These definitions constitute a basis for the elaboration of cooperation schemes between heterogeneous agents (see [14] for an example). However these definitions are very general and we have tried to adapt them to the HRI context.

The main problem in the design of an architecture for HRI is the representation of humans, whose behavior depends on a great number of factors more or less controllable. An idea in this way is the representation of a human agent by a *proxy*.

This approach has been explored and implemented in STEAM and more recently in Machinetta [25], [21], [22]. The idea is that an agent can be represented by a semi-autonomous piece of software called a proxy. The goal is to permit to agents of various types to coexist and to communicate together and to cooperate if it is necessary. A proxy "discusses" with the agent it represents on the one hand and with all the other proxies on the other hand. A proxy is able to choose when it can take decisions by its own and when it is better to leave the control to the agent it represents.

III. DECISIONAL SYSTEM FRAMEWORK

We envision HRI as illustrated in Fig. 1. Two agents (a human and a robot) share a common space and exchange information through various modalities.

Interaction occurs as a consequence of an explicit request of the human to satisfy a goal or because the robot finds itself in a situation where interaction is useful, if not mandatory, for accomplishing the goal.

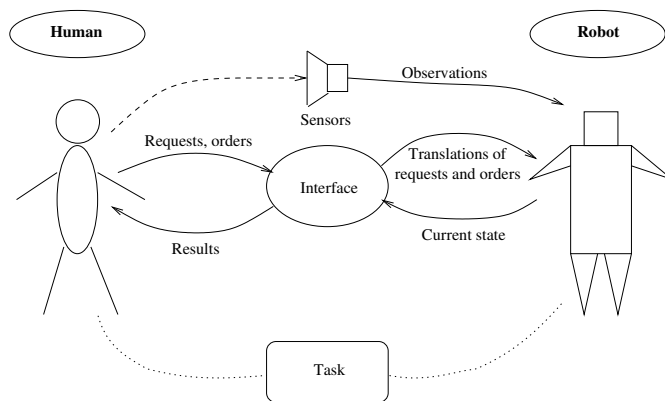


Fig. 1. General description of HRI

In both cases, the robot has a goal to satisfy. An important issue is the notion of engagement, a process in which the robot will have to establish, maintain and end a connection with a human partner.

This will cover goal establishment, selection of an incremental refinement of the task that is intended to satisfy it. The establishment of a connection between the human and the robot will serve the robot to follow human task performance and to monitor the human's commitment to the common goal, and even to influence it.

The HRI we consider in this paper is the common achievement of tasks by two agents - a robot and a human - in order to satisfy a joint goal. The global framework for human-robot interaction at the decisional level of the robot is illustrated in Fig. 2.

Several goals may be sought at a given time, involving possibly several persons. At any moment, there may be several active, inactive and suspended goals. The role of the "Agenda" in this figure is to create and/or abandon goals and to maintain a list of active and suspended goals.

The robot supervision kernel is the central decision-making system of the robot. It is responsible of task selection, refinement and execution. It maintains an integrated view of all robot activities and ensures a global coherence of robot behavior. It is the only entity that can send execution requests to the sensory-motor capacities in the functional level.

The humans encountered by the robot are represented by entities called "InterAction Agents" (IAAs). An IAA is created dynamically and maintained by the "IAA Manager".

The set of active goals entails the incremental execution of a set of tasks in interaction with humans. Each task corresponding to an active or a suspended goal is represented by an entity called "Task Delegate" that is in charge of monitoring the progress towards the goals of both the robot and the IAA and to assess the level of commitment of the associated person. To do so, it controls a set of "Observers" (OBs).

The next section provides a more detailed explanation of

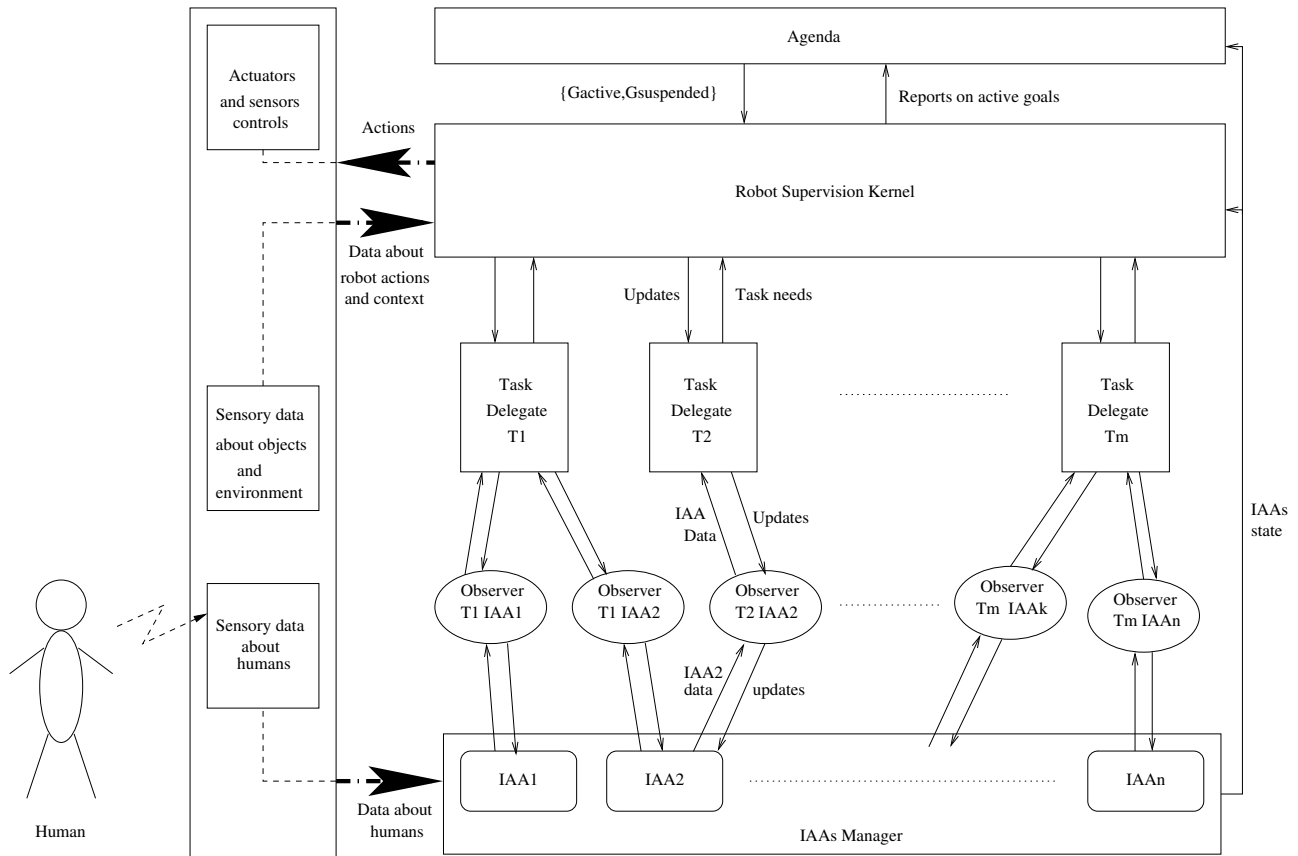


Fig. 2. Decisional framework for a HRI-enabled robot

these entities through an illustrative example.

IV. AN ILLUSTRATED DESCRIPTION

A. Rackham

Rackham[8] is a tour-guide robot (Fig. 3) based on a B21R. Besides improving robustness and efficiency in the robot basic navigation abilities in a dynamic environment, our focus in developing it was to study and test a methodology to integrate HRI abilities in a systematic way.

To test and validate our developments, we have decided to bring regularly our robot to a museum in Toulouse. By regularly, we mean two weeks every three months. Rackham has already been used at the exhibition for hundreds of hours (May 2004, July 2004, February 2005, April 2005), accumulating valuable data and information for future enhancements. The project is conducted so as to incrementally improve the robot functional and decisional capabilities based on the observation of the interaction between the public and the robot.

A number of features have been installed for HRI:

- the detection of dynamic obstacles (i.e., people),
- a vision-based face detector[7],
- a 3D animated head with speech synthesis[4],

- displays and inputs from the touch screen,
- control of robots lights.

All HRIs currently running on Rackham have been classically encoded as event-driven automata with no explicit management of the interactions and no reasoning on human behavior. The next step is to implement the proposed framework for HRI.

B. Rackham's desired capabilities

We show next how the proposed framework will provide a suitable environment to implement flexible human-robot interactions.

Here are some examples of the desired abilities:

- when left alone, Rackham should seek for people to interact with.
- Rackham should be able to detect various types of persons and adapt its behavior to them.
- Rackham should be able to manage two or more interactions, involving several persons, in parallel and for different tasks.
- Rackham should be able to measure the level of commitment of its human interactors and should react accordingly; for instance, detecting that the guided person

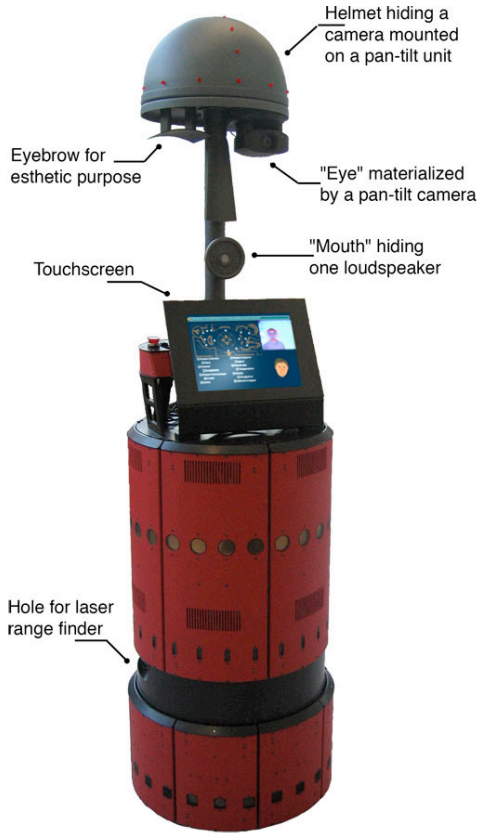


Fig. 3. Rackham: a tour guide robot

follows slowly or is no more interested in the tour he/she has asked for.

C. The management of IAAs

Depending on the context, when the robot detects a human in its vicinity, the IAA Manager creates and instantiates an IAA.

An IAA is defined by its type, its state parameters, its abilities in terms of actions it can perform. Some parameters can be directly perceived by the robot sensors. For each IAA type, a set of “observers” are defined in order to interpret their activity relatively to a given task.

In our example, three types of IAA can be defined:

- **passer-by**: a *passer-by* is a person visiting the exhibition without any interest for Rackham. Its action set is: *move* and *touch the screen*.
- **visitor**: a *visitor* is a person actually guided by Rackham. Its action set is: *move*, *touch the screen* and *follow*
- **troublemaker**: a “*troublemaker*” is a person that blocks Rackham and who does not want to move. Its action set is empty.

The IAA Manager is in charge of creating, deleting and updating the IAAs.

D. The Agenda

The Agenda manages the current set of robot goals. It ensures the consistency between active goals, and determines their priorities, and their causal links.

Based on data provided by the Supervision Kernel, the Agenda determines the relevance of goals and decides to create, suspend, resume or abandon a goal. When a goal is created, it may be associated to the robot alone or to a “team” of agents (the robot and an IAA).

In our example, Rackham can generate goals for *re-charging*, *search_for_an_IIA*, *guide* and *free_the_path*. The two latter goals are involve not only the robot but also an IAA.

E. Robot Supervision Kernel

The Robot Supervision Kernel is responsible of creation, management and execution of plans dedicated to the satisfaction of active goals.

For each new active goal the Robot Supervision Kernel creates a Task Delegate, selects or elaborates a plan and allocates the roles of each team member. Notice that the creation of the Task Delegate is combined with the creation of an Observer (OB) for each human involved in the task performance.

For all the other active goals, the Robot Supervision Kernel has already a plan and is in charge of the execution of the robot part. Whenever an elementary action is performed, the Robot Supervision Kernel forwards this information to all active Tasks Delegates.

Depending on the context, the planning process can be more or less elaborated. Indeed, the presence of humans in the environment raises new issues in motion, manipulation and task planning. We are developing, in coherence with the architecture presented here, a “motion planner in the presence of humans” [24] that can be used not only to plan safe robot paths, but also to plan good, socially acceptable and legible paths, and a high-level “human aware” task planner [18] that is able to deal with constraints imposed by the presence of humans, their needs and preferences.

In our example, the plans are simply obtained by selecting scripts (see Fig. 4).

F. Task Delegates

A Task Delegate represents an unfinished task. It communicates with the Robot Supervision Kernel on the one hand and IAA activity observers (OBs) on the other hand. The role of the Task Delegate is to give to the robot the possibility to follow the course of each task it is involved in.

Fig. 5 exhibits the internal structure of a Task Delegate.

The Task Delegate function is double: to make sure that the task makes progress towards its goal and to assess the level of motivation of the involved humans. We have defined

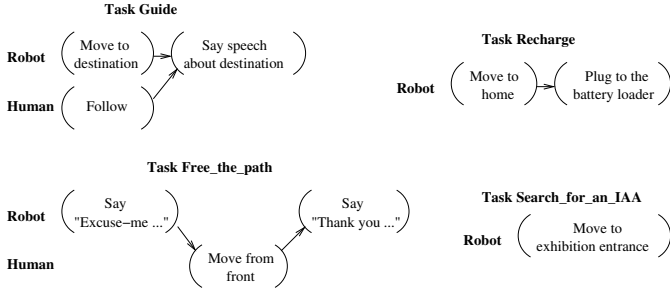


Fig. 4. Scripts for Rackham goals

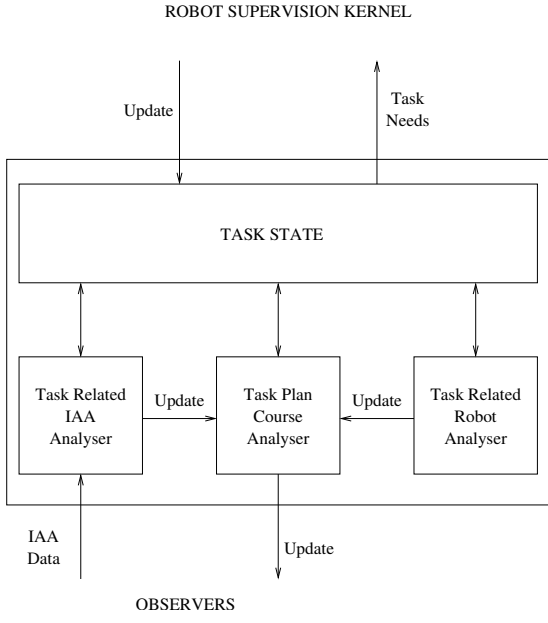


Fig. 5. Structure of a Task Delegate

several phases for a task (PRE-TASK, IN-TASK, CHECK-TASK, POST-TASK, SUSPENDED-TASK) corresponding to its establishment, performance and end.

To illustrate these phases we consider a case where the robot proposes a task to a unique human represented by an IAA called *H*. The robot is called *I* because it is the task instigator. Figure 6 illustrates the transitions from one phase to another.

1) **PRE-TASK**: It is the phase in which the team members decide to participate to the task and define a plan to achieve it: humans must be informed that they have a role to play in the task and must give their agreement (or disagreement) about the way that the task will be performed.

2) **IN-TASK**: It is the core phase of the task, it corresponds to the execution of the task. This phase checks:

- If the task is going smoothly.
- If the human maintains her/his motivation and commitment to do the task, (this is obviously limited to the observation capabilities of the robot).

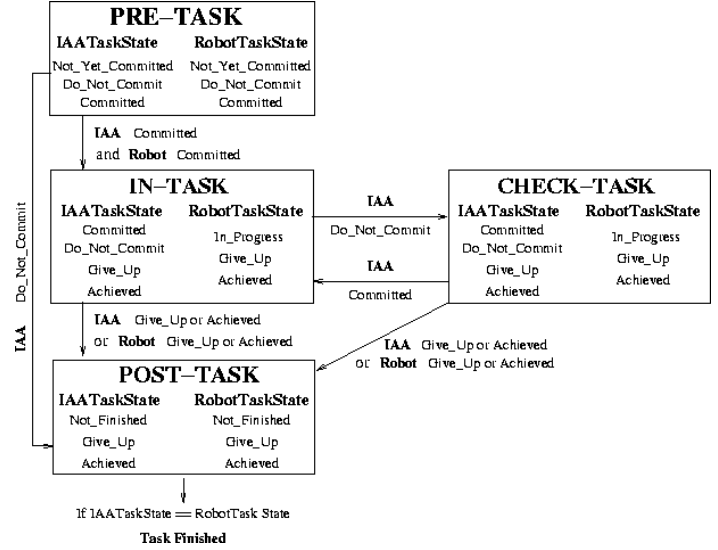


Fig. 6. Task Delegate states

Regarding the progress of the task, the Task Delegate will update the observation rules of OBs and data stored in the Plan Course Analyser thus permitting synchronization between team members. Given the task course and the context, the Task Delegate can decide to stay in IN-TASK or to switch to CHECK-TASK or POST-TASK.

3) **CHECK-TASK**: It is the phase in which, given the observations made, a withdrawal is suspected: so in some conditions depending of the task, things cannot go on. The decision to go in CHECK-TASK is made by the Task Delegate, so it must have the knowledge to be able to do that. Given the context and the result of possible exchanges between the robot and IAA, the Task Delegate can decide to stay in CHECK-TASK or to switch to IN-TASK or POST-TASK.

4) **POST-TASK**: It is the last phase that is reached when the task has been achieved or when it has to be abandoned. A mutual belief about the end of the task must be reached, i.e. the robot must make sure that every agent involved in the task has been informed.

5) **SUSPENDED-TASK**: It is the phase reached when a suspend request comes from the Supervision Kernel. The task has to be suspended and the involved humans should be informed. However, the contact with the associated IAA should be maintained until the task is resumed or abandoned.

G. Observers

An Observer makes the interface between an IAA and a Task Delegate. Its role is to monitor the IAA activity relatively to a task. It is also in charge of detecting situations where the commitment of a team member in a task can be questioned. This will be essentially based on perception and interpretation of human behavior relatively to the task.

While human activity observation and interpretation in a large sense seems out of reach today, the observer will restrict its activity to the extraction of information and the detection of situations that are linked to the task and its context. For example, in the case of a “guide” task, a relevant measure can be the distance between the robot and the guided person as well as its attention to the robot.

V. CONCLUSION AND FUTURE WORK

In this paper we have presented a decisional framework designed for robots operating in a human environment. The particularity of our architecture is that humans are explicitly taken into account.

Our objective was to provide a management of human interaction that can be seen as an integral part of a general robot control architecture. This was done in order to provide a principled way to deal with HRI for task achievement in presence of humans or in synergy with humans.

We also intend to use this approach as a framework in which we will develop and experiment various task planners and interaction schemes. Indeed, the robot will have to plan for itself and for the human in order not only (1) to assess the feasibility of the task (at a certain level) before performing it, but also (2) to share the load between the robot and the human and (3) to explain/illustrate a possible course of action.

Besides, the robot should be able to devise plans that allow it not only to execute its actions but to place itself such as it can be seen by humans or it can observe human activity if it linked to the tasks at hand.

REFERENCES

- [1] J. Allen, G. Ferguson. “Human-Machine Collaborative Planning”, *In Proceedings of the Third International NASA Workshop on Planning and Scheduling for Space*, 2002.
- [2] R. Alami, R. Chatila, S. Fleury, M. Ghallab, F. Ingrand, “An architecture for autonomy”, *International Journal of Robotic Research*, Vol.17, N°4, pp.315-337, Avril 1998.
- [3] R. Alami, T. Simeon, and K. Madhava Krishna. “On the influence of sensor capacities and environment dynamics onto collision-free motion plans.” *Proc. in IEEE Int. Conf. on Intelligent Robots and Systems*, 2002.
- [4] G. Bailly, M. Bîçîşar, F. Elisei, and M. Odisio, “Audiovisual speech synthesis”, *International Journal of Speech Technology*, 6:331-346, 2003.
- [5] C. Breazeal, G. Hoffman, and A. Lockerd, “Teaching and Working with Robots as a Collaboration,” *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2004.
- [6] C. Breazeal, “A Motivational System for Regulating Human-Robot Interaction,” *Proceedings of National Conference of American Association for Artificial Intelligence (AAAI)*, 1998.
- [7] L. Brethes, P. Menezes, F. Lerasle, and J. Hayet, “Face tracking and hand gesture recognition for human robot interaction,” *International Conference on Robotics and Automation (ICRA)*, New Orleans, pp. 1901-1906, 2004.
- [8] A. Clodic, S. Fleury, R. Alami, M. Herrb, R. Chatila, “Supervision and Interaction: Analysis from an Autonomous Tour-guide Robot Deployment,” *International Conference on Advanced Robotic (ICAR)*, 2005.
- [9] P. Cohen and H. Levesque, “Intention is Choice with Commitment,” *Artificial Intelligence*, vol. 42, pp. 213-261, 1990.
- [10] P. R. Cohen, and H. J. Levesque. “Teamwork”. *Nous*, 25(4): 487-512, 1991.
- [11] G. A. Fink, J. Fritsch, S. Hohenner, M. Kleinhagenbrock, S. Lang, and G. Sagerer. “Towards multi-modal interaction with a mobile robot”. *Pattern Recognition and Image Analysis*, 14(2):173-184, 2004.
- [12] T. Fong, I. Nourbakhsh, K. Dautenhahn. “A survey of socially interactive robots,” *Robotics and Autonomous Systems, Special issue on Socially Interactive Robots*, 42 (3-4), pp 143-166, 2003.
- [13] T.W. Fong, C. Thorpe, and C. Baur, “Collaboration, Dialogue, and Human-Robot Interaction,” *Proceedings of the 10th International Symposium of Robotics Research*, Lorne, Victoria, Australia, Springer-Verlag, London, November, 2001.
- [14] N. Jennings, “Specification and Implementation of a Belief- Desire-Joint-Intention Architecture for Collaborative Problem Solving,” *Int. Journal of Intelligent and Cooperative Information Systems*, vol. 2, no. 3, pp. 289-318, 1993.
- [15] K. Kawamura and P. Nilas and K. Muguruma and J. A. Adams and C. Zhou, “An Agent-Based Architecture for an Adaptive Human-Robot Interface”, *Hawaii International Conference on System Sciences (HICSS’03)*, 2003.
- [16] D. Kulic and E. Croft. “Safe planning for human-robot interaction”, *Proc. in IEEE Int. Conf. on Robotics & Automation, New Orleans, LA*, 2004.
- [17] H. J. Levesque, P. R. Cohen, and J. H. T. Nunes. “On acting together.” *In Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, pages 94-99, 1990.
- [18] V. Montreuil, R. Alami “Report on Paradigms for Decisional Interaction” *Cogniron Deliverable D6.2.1, LAAS Internal Report*, January 2005.
- [19] H. Nakajima, et al, “Toward an Actualization of Social Intelligence in Human and Robot Collaborative Systems,” *International Conference on Intelligent Robots and Systems (IROS)*, 2004.
- [20] S. Nonaka, K. Inoue, T. Arai and Y. Mae, “Evaluation of Human Sense of Security for Coexisting Robots using Virtual Reality,” *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2004.
- [21] P. Scerri, et al, “A Prototype Infrastructure for Distributed Robot-Agent-Person Teams,” *The Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2003.
- [22] P. Scerri, et al, “Team Oriented Programming and Proxy Agents: The Next Generation,” *Proceedings of the First International Workshop on Programming Multiagent Systems*, (2004).
- [23] C. Sidner, C. Lee, C. Kidd, N. Lesh. “Explorations in Engagement for Humans and Robots”, *IEEE RA/SJ International Conference on Humanoid Robots*, November 2004.
- [24] A. Sisbot, R. Alami “Architecture of the placement planner for human-robot close interaction” *Cogniron Deliverable D3.5.1, LAAS Internal Report*, January 2005.
- [25] M. Tambe, “Agent Architectures for Flexible, Practical Teamwork,” *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 1997.

What is a Robot Companion – Friend, Assistant or Butler?*

Kerstin Dautenhahn, Sarah Woods, Christina Kaouri, Michael L. Walters, Kheng Lee Koay, Iain Werry

*Computer Science Department
Adaptive Systems Research Group
University of Hertfordshire
College Lane, Hatfield, AL10 9AB, U.K.*

{K.Dautenhahn; S.N.Woods; M.L.Walters; K.L.Koay, I.P.Werry}@herts.ac.uk

Abstract – The study presented in this paper explored people’s perceptions and attitudes towards the idea of a future robot companion for the home. A *human-centred* approach was adopted using questionnaires and human-robot interaction trials to derive data from 28 adults. Results indicated that a large proportion of participants were in favour of a robot companion and saw the potential role as being an assistant, machine or servant. Few wanted a robot companion to be a friend. Household tasks were preferred to child/animal care tasks. Humanlike communication was desirable for a robot companion, whereas humanlike behaviour and appearance were less essential. Results are discussed in relation to future research directions for the development of robot companions.

Index Terms – robot companion, robot-human interaction, social robotics, robot appearance, human perception and attitudes

I. INTRODUCTION

An autonomous robot companion might be viewed as a special kind of service robot that is specifically designed for personal use at home. Robot companions are expected to communicate with non-experts in a natural and intuitive way. Robots designed for the home are a growing industry from both a research and commercial perspective. A survey by the United Nations for example has reported that “robots are set to become increasingly familiar companions in the home by 2007.” By 2007 it is predicted that there will be almost 2.5 million entertainment and “leisure” robots in homes which compares to 137,000 currently [1].

Human-robot interaction research is still relatively new in comparison to traditional service robotics where e.g. robots deliver hospital meals or provide security services, application domains that require relatively minimal human-robot interaction [2]. However, increasingly robots are meant to engage in social-human interaction including e.g. [3], [4], [5], [6]. Robot companions in the home should ideally be able to perform a wide array of tasks including educational functions, home security, diary duties, entertainment and message delivery services, etc. Currently, there are no robots that are able to perform a combination of these tasks efficiently, accurately and robustly. However, research is under way for

developing such robots, e.g. [7]. More and more studies are investigating people’s attitudes towards and perceptions of robots. For example, the Sony Aibo [3], an autonomous entertainment robot for the home designed to elicit emotions and show instincts, learning and growth abilities is often used in child-robot interaction studies. Aibo’s design has been inspired by dog behaviour and appearance [8]. Research by Friedman et al. [9], and Kahn et al. [10] using unstructured play sessions for children and online discussion forums for adults, demonstrated that AIBO was psychologically engaging for both adults and children in terms of life-like essences, mental states and social rapport. However, participants rarely attributed moral standing to AIBO.

Pransky [11] has provided an interesting perspective for the different profiles a future robot companion could take providing the advantages and weaknesses of such a future companion. The ‘Robotic Nanny’ would on the one hand play with children and feed them but on the other hand could lead to a child not having any human interaction and viewing robot interaction as the ‘norm’. A ‘Robotic Assistant/homework companion’ would be able to organise your meetings and research, and track documents, but could lead to the feeling that robot interaction is easier than human interaction. Finally, the ‘Robotic Butler/Maid’ could do all the housework, but may well cause relationship difficulties at home by being too efficient and making one feel redundant.

Investigating the design space of robots is a challenging task that needs to consider various factors [12]. For example, Goetz et al. [13, 14] revealed that people expect a robot to look and act appropriately for different tasks. A robot that performs in a playful manner is preferred for a fun carefree game, but a serious robot is preferred for a serious health related exercise regime. It seems that if a robot cannot comply with the user’s expectations, they will be disappointed and unengaged with the robot. If a robot closely resembles a human in appearance but then does not behave like one, there is the danger of the human-robot interaction breaking down. It could even lead to feelings of revulsion against the robot as in the ‘Uncanny Valley’ proposed by Masahiro Mori [15, 16].

Methods used in psychology can provide a useful starting point for exploring a *human-centred* approach for a robot

* The work described in this paper was conducted within the EU Integrated Project COGNIRON (“The Cognitive Robot Companion”) and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

companion [16]. The studies by Khan [17] and Scopelliti, Giuliani, D’Amico and Fornara [18] are among the first to have used a psychological design framework using questionnaires to explore adults attitudes towards the design of a domestic robot. Khan [17] examined adults’ attitudes towards an intelligent service robot, using a survey which included a variety of different concepts including what people thought robots should look like, how robots could be used for service purposes in the household, how the robots should behave, and how humans have conceived their ideas and images of robots. The survey revealed that most participants were positive towards the idea of an intelligent service robot. Scopelliti et al. [18] investigated people’s representation of domestic robots across three different generations, taking into account gender and educational level, in an attempt to bridge the gap between technological capabilities and user expectations. Their results demonstrated that young people tend to have positive feelings towards domestic robots, whereas elderly people were more frightened of the prospect of a robot in the home.

The European project Cogniron (Cognitive Robot Companion, <http://www.cogniron.org/>) aims to study many of the above topics surrounding the development of robot-human interaction. One of the key aims of the project is to explore what having a robot companion in the home means to people. We adopted a *human-centred* approach to investigate what people’s perceptions and desires for a robot companion are and explored the following research questions:

- Are people accepting of the idea of robot companions in the home?
- What are people’s perceptions of a future robot companion?
 - o What specific tasks do people want a robot companion to perform?
 - o What appearance should a robot companion have?
 - o What are peoples’ attitudes towards a socially interactive robot in terms of robot behaviour and character traits?
 - o What aspects of social robot-interaction do people find the most and least acceptable?

II. METHOD

The part of the methodology reported here is taken from a larger study where subjects participated in a human-robot interaction study within a simulated living room. The interaction trials and their analysis were the main purpose of the larger study, research questions and results will be reported in forthcoming papers. This paper’s subject is the analysis and interpretation of questionnaire data regarding people’s perceptions and attitudes towards a robot.

Design: A series of questionnaires were collected before and after an interaction session with a PeopleBot™ robot. These robots are used as a research platform for research into *future* robot companions by several research partners in the Cogniron project. It is a human-sized, non-humanoid robot of

rather ‘mechanical’ appearance, specifically designed for HRI experiments. The questionnaires relevant to this part of the study were the Cogniron Introductory Questionnaire (providing demographic details), and the Cogniron Final Questionnaire (investigating people’s attitudes and perceptions towards robots).

Sample: Table I illustrates the sample characteristics.

TABLE I: Sample Characteristics

Sample Characteristics (N: 28): Recruited from University of Hertfordshire	
Gender	
Male	50%
Female	50%
Age	
<25	7%
26-35	43%
36-45	29%
46-55	11%
Occupation	
Student	39%
Academic/faculty staff	43%
Researcher	18%
Educational/career background	
Technology related	50%
Non technological (e.g. law)	50%

Instruments:

Cogniron Introductory Questionnaire: This questionnaire enquired about participants’ personal details (age, gender, occupation), level of familiarity with robots, prior experience with robots (at work, as toys, in movies/books, in TV shows, in museums or in schools), and level of technical knowledge of robots were rated according to a 5-point Likert scale.

Cogniron Final Questionnaire: Table II outlines the content of this questionnaire.

Procedure: All subjects completed consent forms before completing the Cogniron Questionnaires and robot interaction trials. After completing the introductory questionnaire, subjects were exposed to a series of human-robot interaction trials, cf. fig. 5. Finally, participants completed the Cogniron Final Questionnaire that enquired about future home robot companions and qualitative aspects of the robot interaction trials.

III. RESULTS

Acceptance of robot companions: Responses for acceptance of computers and computer related technology in the home were more positive compared to responses for the likeability of having a robot companion in the home. 82% of subjects liked or liked very much the concept of computing technology in the home compared to just under 40% for a robot companion (Fig. 1 for likeability of a robot companion in the home). No significant differences were found for gender, age or level of expertise with technology.

TABLE II Content of the Cogniron Final Questionnaire

Section 1 (rated using 5-point likert scale)
<ul style="list-style-type: none"> • What is a robot companion? • Do you like the idea of having a robot companion at home? • What role do you think a future robot companion in the home should have? • What tasks would you like a future robot to be able to carry out? • How controllable, predictable & considerate should a future robot be? • How human-like should the robot appear, behave & communicate? • What speed should a robot companion approach? • How close should the robot come to you? • Should the robot pay attention to what you are doing? • Should the robot be polite and give way if people encounter it? • Should the robot try to find out if you need help before it helps?
Section 2 (rated using 5-point likert scale)
<ul style="list-style-type: none"> • Questions about the subjects' feelings after robot interaction session • Open-ended question about what participants found the most interesting & most annoying during the robot trials. • Open-ended question about whether anything should be changed regarding the robot (appearance, speech, behavior).

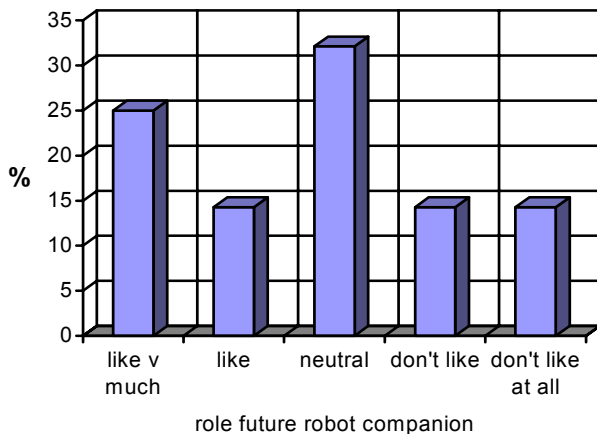


Fig. 1 Likeability of robot companion in the home

The potential role of a robot companion: When asked what role they thought a future 'robot companion in the home should have', the majority of participants wanted the robot as an assistant (79%), a machine/appliance (71%) followed by a servant (46%) (Fig. 2). Fewer people wanted the robot companion as a 'friend' or a 'mate'.[†] Younger subjects suggested that they would like to have a future robot companion in the home as a friend, compared to none of the older subjects ($t(26) = 2.69, p = .01$). No significant

differences were found for gender, or level of expertise with technology.

Task performance for a robot companion: When they were asked what tasks they would like this future robot to be able to carry out, the majority of the subjects wanted the robot to be able to do household (vacuuming) jobs (96.4%).

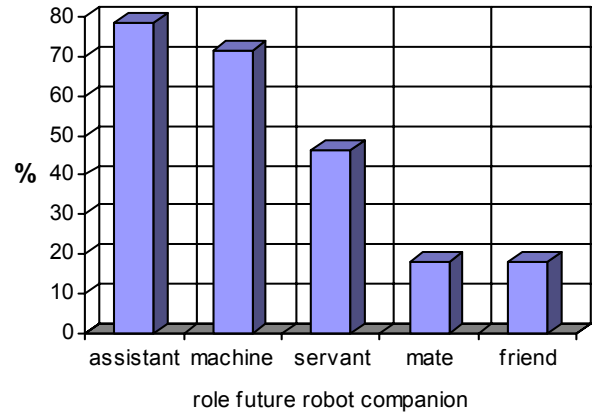


Fig. 2 Desired roles for a future robot companion

Only 10.7% of subjects wanted the robot to be able to look after their children. Guarding the house, entertainment and gardening were also popular choices for robot roles around the home (Fig. 3). No significant differences were found for gender, age or level of expertise with technology related disciplines.

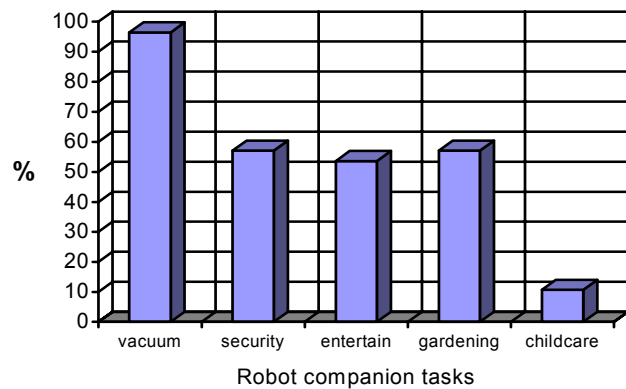


Fig. 3 Preferred tasks for robot companion in the home

Robot companion behaviour traits: Most participants expressed that they would want the behaviour of a robot companion to be highly predictable (54%) or predictable (36%). Only 11% were neutral about the potential predictability of the robot's behaviour. In line with requirements for predictable behaviour, 71% of subjects responded that they would want a robot companion to be highly controllable or controllable. Only one person (4%) stated that they robot should not be controllable. The expression of highly considerate behaviour by a robot companion was also desired by most subjects (86%), and 14% wanted the robot to be behave in considerate manner towards them and other family members. No significant differences were found for gender or level of expertise with technology

[†] Items in fig. 2 & 3 were scored dichotomously as yes/no answers

related disciplines compared to non-technology related subjects.

Robot companion movement: When asked about what speed a considerate robot should approach, the majority (56%) responded neither fast nor slow[‡]. 22% would want the robot to move at a very slow or slow speed. Regarding how close the robot should come to them, 63% said that it should come close to them. Only 4% wanted the robot to come very close to them. The majority of subjects stated that a considerate robot companion should pay attention to what they are doing (85%). Only 15% did not want the robot to pay them any attention. Most subjects noted that a considerate robot companion should be polite and give way to them (70%). With regards to the robot finding out if the subject would want help with a task, 37% of subjects stated that they would prefer the robot to try to find out if they needed help and 41% would want a robot to quietly wait to find out if they needed help. No significant gender or age differences were revealed. However, subjects who had no experience with robots wanted the robot to pay more attention to them compared to those subjects who had experience with robots ($t(25) = 2.41, p = 0.02$). Having no experience with robots was also related to how considerate a robot should be and the wish for the robot to find out if they wanted help more than those who had experience with robots ($t(25) = 2.33, p = 0.03$).

Desired appearance for a robot companion: Participants' responses about human-like appearance, behaviour and mode of communication for a robot companion were somewhat mixed. 71% of subjects would want a robot companion to communicate in a very human-like or human-like manner. However, human-like behaviour and appearance were less desirable. 36% thought that the robot should behave either very human-like or human like, and 29% stated that a robot in the home should appear human-like or very-human like (Fig. 4). Figure 5 illustrates a subject interacting with the PeopleBot™ robot.



Fig. 5 Subject interacting with PeopleBot™ robot in simulated living room

IV. DISCUSSION

The major aims of this study were to explore peoples' attitudes and perceptions towards the idea of having a future robot companion in the home. More specifically, we asked subjects about their level of acceptance of a robot companion in the home, and what types of roles and tasks they would envisage a robot in the home performing.

A summary of the main results indicate that:

- 40% of participants in the current study were in favour of the idea of having a robot companion in the home. This compared to 80% who stated that they liked having computer technology in the home.
- Most subjects saw the potential role of a robot companion in the home as being an assistant, machine or servant. Few were open to idea of having a robot as a friend or mate.
- In terms of specific tasks for a robot companion, 90% stated that it would be useful for the robot to do the vacuuming. This compared to only 10% who would want the robot to assist with child-care duties.
- A future robot companion would need to be predictable, controllable, considerable and polite (possibly based on the current responses of the PeopleBot™).
- Human-like communication was desired for a robot companion. Human-like behaviour and appearance were less important.

It was an encouraging finding that 40% of subjects were in favour of the idea of robot companions in the home, although this figure was lower than the 80% who enjoyed having computer technology in the home. This figure was lower than the 70% of respondents who reported it was a positive idea to have a service robot in the home in the pilot study conducted by Khan [17]. However, this study was different to the current one as it was based on static images of robots and did not involve live interactions. A possible reason for the differences in liking a robot in the home and computer technology could relate to habituation and familiarity effects as computer technology is far more prevalent and accessible to the general public compared to robot companion technology. Also, few subjects were completely against the idea of robot companions

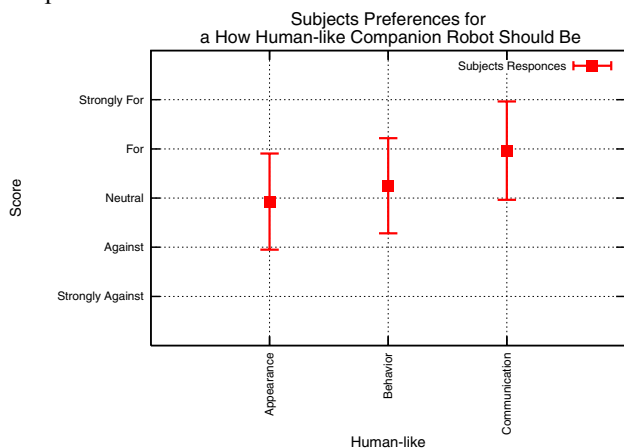


Fig. 4 Subject preferences for how human-like a robot companion should be. Means and standard deviations are shown

[‡] We are aware that “fast” & “slow” are subjective terms. Since this questionnaire was given after the robot interactions, judgements were likely to have been based on the PeopleBot™ speeds which ranged from 200 mm/s – 800 mm/s.

and most subjects appeared to enjoy interacting with the robot trials. Therefore, it is possible that the subjects felt uncomfortable with the idea of a robot companion rather than the reality of the interaction and as robot technology becomes more widely available the differences may become smaller. Dario et al. [19] reported that motor-disabled people were accepting and favourable towards a personal assistance robot in the home, in terms of robot appearance, helpfulness and behaviour. This is in line with our findings although assistive robots are different from more general-purpose robot companions that are relevant to our study.

When the subjects were questioned about the future roles and behaviours of robots in the home, a clear divide emerged. All of the roles which are already, traditionally associated with robots were selected as well as household assistant, gardener and security guard. More than fifty percent of the subjects selected these as roles to be performed in the future. However, roles such as looking after children, being a friend or being a mate were selected by less than eighteen percent of the group. These are all roles which are considered within the 'human domain' and which only a human is able to perform. Some individuals however, in other study can foresee a caretaking role for the robot, for example "I would like having a robot to help me to do something, like taking care of my baby in case he fell from bed" cf. [1]. This could relate to people's perceptions that robots do not possess humanlike personality or character traits. The human fear is sometimes held in the extreme case, that robots could take over the world and replace human abilities [11]. There are no robots currently available on the market able to fully perform the functions of being a child-care assistant or friend, comparable to a human. They are also roles which are the most difficult to prescribe specific actions to in advance, or to specify precisely. These findings are also echoed in a pilot study carried out by Khan [17], which reported that respondents most wanted the robot to perform cleaning tasks. The least cited tasks they wanted the robot to perform were baby sitting, cat/dog watching and reading aloud.

Throughout the study, no differences due to gender or level of technological experience were uncovered for perceptions and attitudes towards a robot companion. Only one age difference was revealed where some younger participants were more in favour of a robot companion taking the role of a friend, compared to none of the older subjects. Likewise, the study by Scopelliti et al. [18] did not find any gender or educational differences towards the idea of robots in the home. However, they found that elderly participants were the most frightened of the prospect of having a robot in the home and showed an element of distrust towards the concept.

Most subjects wanted a robot companion to be controllable and predictable. On one level, any technology for the home should be controllable, in that the user should be able to instruct the device to perform requested actions. However, at the same time, any device should not necessarily require constant supervision, or it ceases to be an aid and instead becomes at best an interface to a task, and at worst something

which slows the user down. This finding could again relate to the unfamiliarity aspect and the possible difficulties in imagining the precise functions of a robot companion in the home. Most people want to be able to understand the logic behind technological devices; therefore it was not surprising that a predictable robot companion was desired. Khan [17] reported similar findings as people in their study did not want the robot to be too smart, but able to conduct limited actions according to its programs.

The fact that subjects wanted a robot companion to have humanlike communication was not a surprising one, as it is a natural human instinct to want to communicate using speech and gestures that are recognisable by humans. Somewhat surprising was that participants did not want the robot to behave and appear in a purely humanlike fashion. However, there is growing research evidence for the Uncanny Valley theory that as a robot approaches pure humanlike appearance, people generally exhibit discomfort and even revulsion towards it. This could also be true of humanlike behaviour for a robot and relates to a human perception of feeling threatened [15, 16].

The current study was exploratory in nature and has revealed many findings that could be relevant for future research ideas and robot companion design. However, a potential drawback of the study could be the self-selected university sample that was recruited to participate. Future studies should attempt to recruit a more representative population sample. Also, the cultural background of subjects, which was not accessed in the present study, is likely to have a significant impact on people's perception of robots. Moreover, none of the sample was older than 55 years, which means that the views of an elderly population are likely to be under represented in this study. This study also relied on subjects *imagining* the role of a possible robot companion in the home, although they interacted in two different trials with PeopleBot™ robots before completing the final questionnaire. It is unclear to what extent these interactions might have influenced the subjects' opinions. Also, the trials were not run within a 'real' home and the scenarios did not cover all possible types of interactions and tasks that might occur in a home. In future trials we hope to explore different types of tasks in order to allow subjects to gain a greater understanding of possible robot capabilities. We are also aiming to use longitudinal experimental paradigms with the same sample of subjects, which will allow us to measure whether perceptions and attitudes towards a robot companion change over time as familiarity increases. A further aim is to change the behaviour and appearance of the robot to determine how these factors influence people's views long-term.

V. CONCLUSION

To conclude, the current study explored people's perceptions and attitudes towards the idea of a robot companion in the home. Interesting and positive results have emerged, indicating that a large proportion of people are

favourable to the idea of a robot companion. Results have highlighted the specific roles and tasks that people would prefer a robot companion to perform in addition to the desired behavioural and appearance characteristics. The finding that people frequently cited that they would like a future robot to perform the role of a servant is maybe similar to the human 'butler' role. Ogden & Dautenhahn [20] considered the concept of 'robotic etiquette' in relation to body movements and positioning to convey polite interactions to advance the social-interaction abilities of robots. A deeper exploration of the necessary training guidelines to become a competent butler could aid the design of future robot companions. For example, butlers need to know how to wait discreetly until given an order to perform a task, and to know when to speak to their employer. This requires great awareness and sensitivity to social situations. Other tasks that a competent butler should be able to perform include: supervising staff, ensuring safety and security, answering the door/phone, preparing meal services and social events, and valet duties etc. A mixed quantitative and qualitative approach taken was advantageous for this research program.

As a cautionary note, one has to be aware that in any HRI study it is practically and methodologically impossible to control for *all possibly relevant* factors that might influence an experiment, as well as providing a large and balanced sample size and rigorous analysis. Thus, exploratory studies like the one reported in this paper, while they often raise more questions they are able to answer conclusively (generalizable to all possible robot appearances, behaviours, contexts, tasks, human subjects etc.), serve an important role in HRI research: they can provide a starting point for identifying relevant future research directions that then need to be investigated in more depth in focussed studies.

REFERENCES

[1] B. News, "Robots set to get homely by 2007," in <http://news.bbc.co.uk/1/hi/technology/3764142.stm>, 2004.

[2] D. Wilkes, R. Alford, R. Pack, R. Rogers, R. Peters, and K. Kawamura, "Toward socially intelligent service robots," *Applied Artificial Intelligence Journal*, vol. 12, pp. 729-766, 1997.

[3] Sony, "Sony Entertainment Robot Europe <http://www.aibo-europe.com/>," 2004.

[4] C. L. Breazeal, *Designing sociable robots*. The MIT Press, 2002.

[5] L. Canamero, "Playing the emotion game with Felix: What can a LEGO robot tell us about emotion?" in *Socially intelligent agents: Creating relationships with computers and robots*, K. Dautenhahn, A. Bond, L. Canamero, and B. Edmonds, Eds. Massachusetts, Kluwer Academic Publishers, 2002, pp. 69-76.

[6] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu, "Robovie: An interactive humanoid robot," *Industrial robot: An international journal*, vol. 28, pp. 498-503, 2001.

[7] Wakamaru, "Target functions for "Wakamaru" under development," in http://www.compukiss.com/populartopics/retire_govhtm/article923.htm, 2004.

[8] J. Pransky, "AIBO - the no. 1 selling service robot," *Industrial robot: An international journal*, vol. 28, pp. 24-26, 2001.

[9] B. Friedman, P. H. Kahn (Jr.), and J. Hagman, "Hardware companions? What online AIBO discussion forums reveal about the human-robotic relationship," *Digital Sociability*, vol. 5, pp. 273-280, 2003.

[10] P. H. Kahn (Jr.), B. Friedman, D. R. Perez-Granados, and N. G. Freier, "Robotic pets in the lives of preschool children," *Proc. CHI*, 2004.

[11] J. Pransky, "Social adjustments to a robotic future," www.robot.md/publications/sfra.html, pp. 1-10, 2004.

[12] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, pp. 143-166, 2003.

[13] J. Goetz and S. Kiesler, "Cooperation with a robotic assistant," *Proc. CHI'02 Conference on Human Factors in Computing Systems*, New York, USA, pp. 578-579, 2002.

[14] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behaviour to tasks to improve human-robot cooperation," *Proc. IEEE Ro-man*, pp. 55-60, 2003.

[15] K. Dautenhahn, "Design spaces and niche spaces of believable social robots," *Proc. IEEE Ro-man*, 192-197, 2002.

[16] S. Woods, K. Dautenhahn, and J. Schulz, "The design space of robots: Investigating children's views," *Proc. IEEE Ro-man*, pp. 47-52, 2004.

[17] Z. Khan, "Attitude towards intelligent service robots," NADA KTH, Stockholm 1998.

[18] M. Scopelliti, M. V. Giuliani, A. M. D'Amico, and F. Fornara, "If I had a robot at home. Peoples' representation of domestic robots," in *Designing a more inclusive world*, S. Keates, J. Clarkson, P. Langdon, and P. Robinson, Eds. Springer, 2004, pp. 257-266.

[19] P. Dario, E. Gugliemelli, and C. Laschi, "Humanoids and personal robots: Design and experiments," *Journal of Robotic Systems*, vol. 18, pp. 673-690, 2001.

[20] B. Ogden and K. Dautenhahn, "Robotic etiquette: Structured interaction in humans and robots," *Proc. SIRS2000*, The University of Reading, England, pp. 353-361, 2000.

Hey, I'm over here – How can a robot attract people's attention?*

Markus Finke

*Neuroinformatics and Cognitive Robotics Group
Faculty of Informatics and Automatization
Technical University Ilmenau
P.O.Box 100565, 98684 Ilmenau, Germany
Markus.Finke@Stud.Tu-Ilmenau.De*

Kheng Lee Koay, Kerstin Dautenhahn,

Chrystopher L. Nehaniv, Michael L. Walters and Joe Saunders
*Adaptive Systems Research Group
School of Computer Science, University of Hertfordshire
College Lane, Hatfield, Herts AL10 9AB, United Kingdom
K.L.Koay@herts.ac.uk*

Abstract— This paper describes how sonar sensors can be used to recognize human movements. The robot distinguishes objects from humans by assuming that only people move by themselves. Two methods using either rules or Hidden Markov Models are described. The robot classifies different movements to provide a basis for judging if a person is interested in an interaction. A comparison of two experiment results is presented. The use of orienting cues by the robot in response to detected human movement for eliciting interaction is also studied.

Index Terms— Human-Robot Interaction, Social Robotics.

I. INTRODUCTION

Service robots are being developed for applications to assist humans which involve dialogue and/or communication with humans. Environments include offices or department stores, where service robots can potentially provide useful information on products (e.g. Boehme et al. [1], Gross et al. [2]). Should a robot directly approach and verbally address a customer? This behaviour might be acceptable if displayed by a human sales assistant, but might be interpreted as “intrusive” or “pushy” behaviour if performed by a robot. Would it be beneficial to have a robot capable of interesting a person in interaction with it in a more “gentle” way (e.g. by leaving it to the customer to approach the robot and initiate the interaction). Could certain movement cues provided by the robot elicit such self-initiated human behaviour? Therefore we investigated the following research questions: 1) How can a robot detect that a human is interested in interacting with it? 2) Can simple orientation movements be used to encourage a person to interact with a robot?

In order to address these research questions, we developed and experimentally evaluated two computational methods for detecting human movements using sonar sensors on a Peoplebot™ robot. Also we studied in experiments the reaction of human subjects towards the robot in conditions involving orientation cues.

The conventional approach for detecting human movement is normally performed using vision systems [3]. We believe

that in a human-robot interaction scenario, the dynamics of multi-modal interaction are often more important than the precise detection of particular features in the environment. It may also be possible to use sensory fusion for detecting human movement in the future.

The commercially available Peoplebot™ robot was used in this experiment and has various sensors, including infrared, sonar, contact sensors, and an onboard camera. We investigated using a sonar-based movement detection system, since non-vision sensors are widely used with success in the field of mobile robotics, especially in the area of obstacle detection. Buchberger et al. [5] uses a combination of laser and sonar sensors which avoid static and dynamic obstacles by recognizing objects in realtime. Salter et al. [4] using arrays of infrared sensors to detect human behaviour.

Sonar sensors can be error-prone as sometimes the data from a sensor can be lost. Sources of error include: ultrasonic waves not being deflected back directly to the sensor, but to other objects, then back to the robot. Detected distances maybe overestimated and the robot may collide with an object. Crosstalk may occur if more than one source emits ultrasonic waves. The received echoes can be sent by another source, so that the sensor detects an object as closer than it really is. Joerg and Berg [6] describe a method that defines an echo for each sensor so that it can distinguish between its own echo and that coming from any other source. They use pseudo-random sequences to get independent ultrasonic waves. Every sensor can also be used to identify echoes from other sensors. This information can be used for triangulation. Finally all sensors can emit ultrasonic waves at the same time so that obstacles can be detected earlier.

The basic research approach is presented in section II. In section III we describe two algorithms for the recognition of human movements. The first one is rule-based and the second one uses Hidden Markov Models (HMMs). A comparison of both methods is shown in section IV. Section V provides an analysis of how human behaviour may be related to the robot's behaviour.

II. BASIC RESEARCH APPROACH

Sonar sensors cannot distinguish between an object and a person, and can only give two kinds of data: 1) There is an object at a measured distance. 2) There is no other object

*The work was carried out during an internship of Markus Finke at University of Hertfordshire, hosted by the Adaptive Systems Research Group. The work described in this paper was partially conducted within the EU Integrated Project COGNIRON (“The Cognitive Companion”) funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

between that detected and the robot, because ultrasonic waves cannot go through objects.

The change of data over time is important because movements cause variations at every sensor. For the purpose of this paper we assume that only people can move by themselves and that moving objects detected at a height of one meter are usually associated with a moving person.

We also assume that the robot itself is static. Otherwise its movements cause significant changes of the data and the system cannot know if this is caused by a person or by the robot.

In order for the robot to know that a person is interested or wants to interact we take E.T. Hall’s [7], [8] “social distances” into consideration. At a certain proximity, it could be assumed that a person wants to interact with the robot. The spatial distances between a robot and a human are discussed in Walters et al. [9]. The generally recognized personal space zones between humans (e.g. northern Europeans) are well known and are discussed in Lambert [10]. It is also important to note that we cannot classify every person that has entered the robot’s social zone as interested in interacting with the robot, as the person may be just passing through the area. It is also safe to assume that people that are outside the robot’s social zone are probably not interested in interacting with the robot. Therefore, the system should identify different movements before deciding if the person is interested in interacting with the robot.

III. THE ALGORITHMS

A. A rule-based approach

The algorithm (see Fig. 1(a)) concentrates on the following information: 1) The distance between the robot and the human subject, 2) the duration a human subject spends within the detection window of each sonar sensor, and 3) the initial and the final distances between human subject and robot when the human subject entered and left the detection window, respectively.

The collected data shows that sometimes the signal is lost for 3 or 4 timesteps (0.3 - 0.4 seconds). The received values are set to -1 to indicate the sensor error condition which is then ignored.

Sonar sensors readings are never stable, even in a static environment. This limitation does not preclude its usage as human movements usually cause more significant changes in the sensor readings. We used two different threshold values (i.e. $k1=0.8$ and $k2=1.35$) to assist in identifying human movements. These values were defined based on the ratio of previous and current sensor readings of the distances between the subject and the sensor, and were obtained empirically through trial-and-error. These threshold values are plotted on figure 1(b), where $k1$ and $k2$ each represent the border lines that separate regions B and C, and regions C and A respectively. Different regions correspond to a person entering (zone A) or leaving (zone B) the area of detection of the sensor. If the ratio of previous and current sensor readings is in

zone C, this means either no significant change has occurred or the person is still in the area of detection.

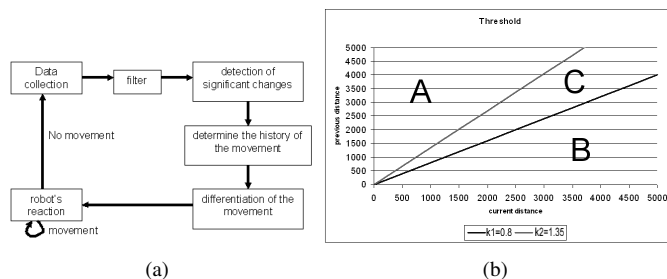


Fig. 1. (a) The main modules of the rule-based algorithm, (b) Thresholds indicating human movements for a single sensor.

Usually only one or two sensors will be involved in detecting a person’s approach behaviour (depending on how far the person is from the sensors and the sensing angle of the sound beams of the sensors). There will be no significant changes in the sensory readings as the person approaches. However by comparing each of the current sensory readings of the involved sensors with the average sensory reading over a period of 10 timesteps (i.e. 1 second), the system can recognise human approach behaviour.

The rows of the matrix in figure 2 each show how a given sensor has been activated over time. If a person approaches the robot, one or two sensors will be activated several times in a short sequence.

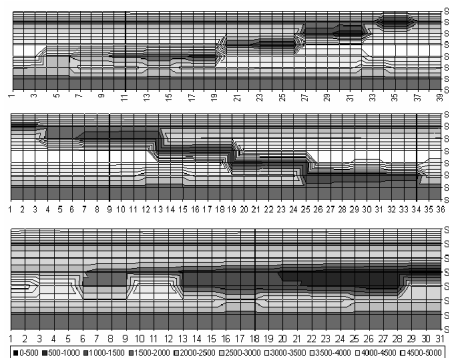


Fig. 2. Patterns of movement. The y-axis displays the eight sensors. S1 points to the left, S8 to the right, S4 and S5 point to the front of the robot. The x-axis displays the time. The closer an object is detected the darker the gradient is. Top: Diagonal movement from the far left corner to the right side of the robot. Middle: Movement straight from the right side of the robot to the left. Bottom: Shows the movement of a person approaching the robot.

For detecting other human movement behaviours, the system will have to look at all the sensory readings over a period of 40 or 50 timesteps. The history of the sensory readings is usually stored in a table, where each column represents the sensory readings at each timestep (see figure 2). The system identifies movement by tracking the movement of the darkest gradient – corresponding to the closest object detected — along the sensor axis (i.e. row) over a period of 40 or 50 timesteps (i.e. columns). Movement usually involved the

darkest gradient moving across 4 rows in a sequential manner over a period of 40 or 50 timesteps (i.e. each of four sonar sensors sequentially detect a person over a period of time). By studying these examples of human movement data recorded from the sensors in such matrices, 11 rules were hand-coded to detect and classify the types of motion.

B. Using Hidden Markov Models

Hidden Markov Models (HMMs) are a technique using finite automata with probabilistic transitions to model the generation of observations corresponding to different patterns of a system's behaviour. The basic ideas are explained in the tutorial by Rabiner [11]. HMMs are widely used for pattern recognition. For example; Billard and Calinon [12] used HMMs to recognize and produce matching behaviours in a skill learning (imitation) context. Westeyn et al. [13] explain a system that detects gestures to control a car radio.

Significant examples are necessary to train the HMMs in order to use them in an application with actual data. Different movements in the environment of a robot cause different patterns in the detected distances over time. Sample data is illustrated in figure 2.

We used the Georgia Tech Gesture Toolkit (GT²k) built on top of an HMM system from Cambridge University – the same system used for the above two applications [13]. With a few modifications GT²k automatically ran the HMM algorithms. The most extensive task was the preparation of training data for the system.

The HMMs are time-invariant, but cannot easily handle cases where the same movement they were trained to detect occurs either at different distances or in different environments. Therefore, we standardized the data from the eight sensors, but lost the ability to distinguish between movements towards or away from the robot. Also, as movements towards the robot were only recognized badly as a result of using a single HMM approach, a two HMM was used.

Preprocessing the data got rid of sequences where the sonar data was lost, and these were replaced by the distance measured in the following timestep. The system receives the data and continuously averages the last ten timesteps. This is stable if nothing happens in the environment of the robot. Otherwise the difference between the current distance and average will be significant.

We used two eight-state HMMs where transitions cannot go back to a previous state, but can stay at the same one, go to the next one or even skip one state (so-called 'right-to-left models'). The first model is responsible for movements close to the robot, so that people who are interested in the robot or want to interact, are detected. The second model recognizes movements from one side of the robot to the other one. This model classifies people who are not interested, or show only a little interest, but move on. These people might be interested if they notice that the robot watches them and maybe turns towards them.

For detecting a movement, the values of the differences between current distance and average of the last seconds

were saved in a text-file. This was repeated for every sensor independently using the first HMM (i.e. approach detection). Only if no movement towards the robot is detected, will the values of all eight sensors together be saved in a second text-file, which was used as input for the second HMM (i.e. left or right movement detection).

C. Behaviour of the robot

The robot behaves the same way for both algorithms depending on the recognized movement. If a person approaches the robot closer than one meter, the robot will assume that this person is interested or wants to interact. In this case the robot turns head-on to the person, because people are used to talking face to face during an interaction. The distance of one meter is chosen with regard to E.T. Hall's "social distances". He subdivides the environment of a person into intimate, personal, social and public zones. The personal area is an adequate distance for human-human interaction and we assumed in this work that it also applies to human-robot interaction.¹

If the robot detects a movement from one side to the other, it will assume that the person is not interested. It is also possible that a person did not realize that the robot was working and watching him. The robot will then turn 45 degrees in the direction the person is moving. This gives feedback to the person that he has been detected. The person might then become interested in the robot and approach. If the person does not come close to the robot but moves on, the robot will turn back to its previous position.

Collection of sonar data was temporarily suspended as soon as the robot started turning. Otherwise, the robot would experience sensory input similar to when a person is moving from left to right, as it turns from right to left, and vice versa. If the robot turns, one sensor will detect distances that its neighboring sensor has detected earlier and there will be a significant variation that would be interpreted as the "detection" of a moving person.²

IV. COMPARISON OF BOTH ALGORITHMS

We carried out two experiments in order to compare both algorithms. The first one took place in the same environment as the training phase. Twelve people, who were not involved in the training, moved around the robot. The second experiment took place in a new environment with people who were neither involved in the training nor in the first experiment. This experiment demonstrated the ability of the algorithms to generalize in a new environment.

¹But compare also the results of [9], showing strong individual differences between people on whether human-human interaction distances are generalized to their interactions with robots. In particular, some persons appear to treat robots more as objects (to which human-human social distances do not apply).

²This limitation of the robotic system is analogous to the fact that humans are blind to changes in visual scenes that occur during saccadic eye movement [14].

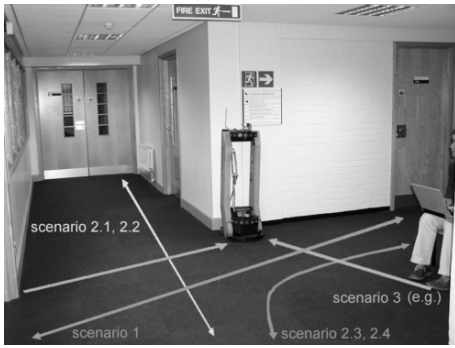


Fig. 3. Environment of the first experiment. The movements of the three scenarios are shown by the arrows. Scenario 1 corresponds to the subject moving from “right to left” and “left to right” in front of the robot. Scenario 2.1 and 2.2 have the subject moving past the robot on its right side from front to back and visa versa. Scenario 2.3 has the subject moving along a curved path from the robot’s front to the passage at the robot’s left. Scenario 2.4 is the same but with the subject moving in the opposite direction. Scenario 3 corresponds to the subject moving “forward” in a straight line from anywhere within a semi-circle ahead of the robot, stopping in front of the robot.

A. Experiment 1

Each person received written instruction for movements subdivided into three scenarios (see fig. 3) before the experiment started. The environment of this experiment is shown in figure 3, where the movements of the three scenarios are indicated by the arrows.

In order to compare both the algorithms using exactly the same human movement data, each algorithm was used for an online test of six of the twelve cases, and the movement data were collected. The data collected during online testing of an algorithm were then later used in offline testing of the other algorithm. Therefore a total of twelve (six online and six offline) human movement cases was tested on each of the algorithm. Note that the movement data recorder stopped recording movement data as soon as an algorithm recognised a movement. Because the HMMs algorithm required more movement data than the rule-based method to classify a movement, it was expected that the HMMs would perform badly during offline testing with data collected from the rule-based method.

The results are shown in table I. The “incorrect” classifications include all examples that were classified incorrectly or were not classified.

Both algorithms were not trained for the movements in scenario 2. Scenarios 2.3 and 2.4 were similar to right→left and left→right training examples, but scenarios 2.1 and 2.2 were completely different from any of the training examples. The result shows that both algorithms could not detect “new” movements very well, especially the HMMs with a correct classification rate of 50% or less.

Table II shows the online and offline test results of right→left, left→right and forward movements. The columns and the rows of the table represents the online and offline test results of the algorithms respectively. This table shows in detail how many of the test examples were detected correctly

by one method, were detected incorrectly by the other method and vice versa.

The results shown in table III indicate that overall the algorithm using HMMs performed better than the rule-based method, but were worse on-line than rule-based offline for the left to right movement.

TABLE I
ONLINE TEST RESULTS OF THE FIRST EXPERIMENT

Movement	Online Test Results					
	HMMs Algorithm			Rule-Based Algorithm		
	Correct (%)	Incorrect (%)	N	Correct (%)	Incorrect (%)	N
Right to Left	78	22	18	70	30	20
Left to Right	72	28	18	60	40	20
Scenario	2.1	67	33	6	50	6
	2.2	33	67	6	50	6
	2.3	50	50	6	100	6
	2.4	33	67	6	83	6
Forward	70	30	20	90	10	21

TABLE II
ONLINE VS. OFFLINE TEST RESULTS OF THE FIRST EXPERIMENT

			Online Test Results									
			HMMs Algorithm			Rule-Based Algorithm						
			Correct	Incorrect	Total	Correct	Incorrect	Total				
Offline Test Results	Rule-Based Algorithm	Right to Left	Correct	10	1	11	HMMs Algorithm	Right to Left	Correct	12	2	14
			Incorrect	4	3	7			Incorrect	2	4	6
			Total	14	4	18			Total	14	6	20
		Left to Right	Correct	12	2	14		Left to Right	Correct	12	6	18
			Incorrect	1	3	4			Incorrect	0	2	2
			Total	13	5	18			Total	12	8	20
	Forward	Correct	7	5	12	Forward	Correct	19	1	20		
		Incorrect	7	1	8		Incorrect	0	1	1		
		Total	14	6	20		Total	19	2	21		

Note: The results of the online and offline tests are shown in the columns and rows of the table respectively.

TABLE III
EXPERIMENT 1 RESULTS SUMMARY

Movement	Online	Offline	Offline	Online	Overall: Online-Offline	
	HMMs (%)	Rule-Based (%)	HMMs (%)	Rule-Based (%)	HMMs (%)	Rule-Based (%)
Right to Left	78	61	70	70	74	66
Left to right	72	78	90	60	81	69
Forward	70	60	95	90	83	75
Total	73	66	85	73	79	70

B. Experiment 2

The second experiment took place in a public corridor (see figure 4). The people were passers-by and were not instructed how to move or to behave in front of the robot, nor were they informed that an experiment was in progress. It should show how well the algorithms work in a different environment, with people who were not involved in the training process. We collected data in five different conditions or ‘rounds’. The difference between the conditions was in the behaviour of the robot. As in Experiment 1, the behaviour of the robot depended on the detected movement. In this experiment, the



Fig. 4. Environment of the second experiment. Sample paths that were chosen by the detected people are shown by the arrows.

robot responded differently in different conditions. During the first round the robot did not react. In the second and fourth round it turned 20 or 50 degrees respectively to the direction the person was moving (i.e. follow direction). In the third and fifth round the robot turned 20 or 50 degrees respectively into the direction the person came from (i.e. away direction).

The first condition, when the robot did not turn, lasted ten minutes. The other four conditions lasted five minutes each. Total numbers per round of observed persons varied and overall we tested 152 subjects. Most of the subjects movements were moving from one side to the other (relative to the robot). The robot stood near the entrance of a main corridor. During the trials, we observed that the majority of the subjects slowed down when they noticed the robot, but most of them moved on while looking at the robot. There was only one subject that became very interested in the robot, and approached it.

“Back” means that people came through a door behind the robot, passed its left side and turned left or passed its right side and turned right. In both cases they did not cross in front of the robot. “Forward” means the corresponding movement towards the door behind of the robot.

The robot used the rule-based algorithm during the experiment to trigger movement when a person was detected. The stored data was later tested offline using the HMMs. The results are shown in table IV.

There were eight cases where two subjects moved from two different directions in front of the robot. In six of these cases, the robot managed to detect only one subject’s movement instead of the two movements. For the other two cases, the robot failed to detect these movements. Note, we did not expect the robot to accurately detect simultaneous movements of more than one subject as both algorithms were built for detecting a single subject movement at a time.

The algorithm using HMMs is better overall than the rule-based one. However, for movements from left to right the rule-based method is better than the HMMs. The HMMs need to be trained with different data to improve the recognition of movements from left to right. The comparison of both algorithms’ performance is shown in figure 5.

TABLE IV
SIMULATION (HMMs) RESULTS OF THE SECOND EXPERIMENT

		Online Test Results of Rule-Based Algorithm										Total			
		Robot in Static		Robot Rotate Follow Direction				Robot Rotate Away Direction							
				20°		50°		20°		50°					
		%	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct		Incorrect		
Offline Test Results of HMMs Algorithm	Subject's Movement	Right to Left	Correct	100	0	60	40	50	42	100	0	60	13	93	
			Incorrect	0	0	0	0	8	0	0	0	13	13	7	
		Left to Right	Correct	60	30	43	29	20	20	25	25	25	25	60	
			Incorrect	10	0	14	14	40	20	50	0	50	0	40	
		Forward	Correct	100	0	-	-	-	-	-	-	-	-	-	100
			Incorrect	0	0	-	-	-	-	-	-	-	-	-	0
	Back	Correct	100	0	100	0	33	33	0	100	0	100	93		
		Incorrect	0	0	0	0	0	33	0	0	0	0	7		
	Total		93	7	72	28	51	49	58	42	49	51			

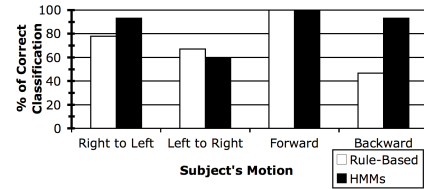


Fig. 5. Experiment 2 Results Summary

V. ANALYSIS OF THE VIDEO DATA

Finally, we analyzed the video data with regard to the behaviour of people when they noticed the robot. A person’s behaviour is grouped under one of five categories according to the robot’s reaction. The total number of each of the five groups differs from the total number of these groups in table IV because sometimes the robot turned the wrong way due to misclassification. The difference in these numbers is also caused by groups of people that were detected by the robot as one person, but not all reacted the same way.

The video data showed the following seven behaviours of people, organized from those that showed the highest interest to those who showed no interest in the robot: **a** - approaches the robot, **b** - stops, speaks (excitedly) then continued along the same path, **c** - stops and watches, then continued along the same path, **d** - watches and slows down, **e** - watches while walking on, **f** - only a short glimpse at the robot while walking on and **g** - ignores robot.

Analysis of human behaviour just by watching it is very difficult because the interpretation is influenced by the personal opinion of the observer. Because of this the video data was interpreted twice by two different people. The level of inter-rater agreement is 84%.

Figure 6 shows the results of this analysis. The surprisingly high ratio of people who ignore the robot was due to some people who moved along the corridor several times, but were not interested in the robot when they came for the second or third time. The ratio was also influenced by some members of our laboratory who already knew the robot and its behaviour.

Table V illustrates subject behaviours in respond to robot’s action. If the robot classified the movement incorrectly, the resulting behaviour of the people was added to the group according to the robot’s action. If the robot did not detect a

movement, the behaviour of the person was classified “doesn’t move” since the robot did not move.

The results with respect to the behaviour of the observed people show that most of them looked at the robot, while walking on. However, the orientation cue of the robot, which was meant to attract the person’s attention (rotation), seemed not to be sufficient for eliciting a response. This may be due to the robot movement detector algorithms taking about 4 or 5 seconds to detect a movement. By the time the robot moves, the person had already passed the robot. Also, the experiments were carried out in a busy university where people often walked past quickly and were not distracted easily. Under these difficult conditions it is probably not surprising that a simple orientation cue did not have any major effect.

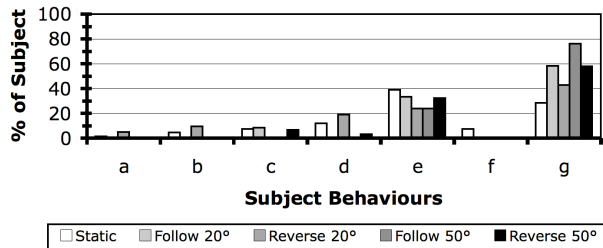


Fig. 6. Subjects’ Reaction to Robot’s Behaviours

TABLE V

VIDEO ANALYSIS OF PEOPLE’S REACTION TO ROBOT BEHAVIOUR

Subject's Reaction	Robot's Reaction				
	Static	Follow 20°	Reverse 20°	Follow 50°	Reverse 50°
a	1	0	1	0	0
b	3	0	2	0	0
c	5*	1	0	0	2
d	8	0	4*	0	1
e	26*	4	5	5	10
f	5	0	0	0	0
g	19	7	9	16	18
Total (N)	67	12	21	21	31

Notes: *2 persons stopped as they wanted to know if they can pass the robot; 2 persons stopped as they talked to each other and then moved to different directions, ♦ once a person appeared shortly after another and talked to him about the robot and ♣ one person looked bemused when the robot turned.

VI. CONCLUSION

We have created two algorithms to detect human movements in the environment of a robot just by using sonar sensors. One algorithm is rule-based and analyzes the sonar data in order to find significant changes over time. The second one uses Hidden Markov Models to recognize a pattern in the data.

Both algorithms were implemented on a PeopleBot™ and their reliability was compared in two experiments. The second experiment also tested the ability to generalize in different environments. The results of both experiments show that both algorithms work adequately, but the one using Hidden Markov Models works better and detects the movements correctly in approximately 80% of the cases. The reliability of the

algorithms can be improved in the future by incorporating different movements which happen in real scenarios.

The detected movements are used by the robot to interpret the behaviour of a person. We assumed that people who are interested in the robot and want to interact, approach the robot. The reaction of the robot to interact with the subject will depend on this interpretation.

With respect to people’s behaviours, we found that most people looked at the robot when the robot is in sight, while walking on. However, the robot’s orientation cue (rotation) was not enough, perhaps due to the robot’s slow reactions. One of the solutions could be first using a voice system to attract the attention of people that have already moved past the robot, then followed by the orientation cue. Future work needs to investigate other robot cues (e.g. movement, speech, gestures) or more likely, a combination of various robot cues that will be able to attract attention and encourage approach and engagement in an interaction with the robot.

REFERENCES

- [1] H.-J. Boehme, T. Wilhelm, J. Key, C. Schauer, Ch. Schroeter, H.-M. Gross and T. Hempel, “An Approach to Multi-modal Human-Machine Interaction for Intelligent Service Robots”, *Robotics and Autonomous Systems*, vol. 44, pp. 83-96, Elsevier Science.
- [2] H.-M. Gross, H.-J. Boehme, J. Key and T. Wilhelm, “The PERSES Project - a Vision-based Interactive Mobile Shopping Assistant”, *Künstliche Intelligenz*, vol. 4, pp. 34-36, 2000.
- [3] W. Zajdel, Z. Zivkovic and B. Kröse, “Keeping track of humans: Have I seen this person before?”, Accepted for *ICRA*, 2005.
- [4] T. Salter, R. te Boekhorst and K. Dautenhahn, “Detecting and analysing children’s play styles with autonomous mobile robots: A case study comparing observational data with sensor readings”, *Proc. IAS-8*, 2004, pp. 61-70.
- [5] M. Buchberger, K.-W. Joerg and E. von Puttkamer, “Laserradar and Sonar Based World Modeling and Motion Control for Fast Obstacle Avoidance of the Autonomous Mobile Robot MOBOT-IV”, *IEEE International Conference on Robotics and Automation*, 1993, pp. 534-540.
- [6] K.-W. Joerg and M. Berg, “First Results in Eliminating Crosstalk and Noise by Applying Pseudo-Random Sequences to Mobile Robot Sonar Sensing”, *Proc. EUROBOT '96*, pp. 40-45.
- [7] E. T. Hall, *The Hidden Dimensions: Man's Use of Space in Public and Private*, The Bodley Head Ltd, London, UK, 1966.
- [8] E. T. Hall, “Proxemics”, *Current Anthropology*, vol. 9, no. 2-3, pp. 83-108, 1968.
- [9] M. L. Walters, K. Dautenhahn, R. te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. Nehaniv, D. Lee and I. Werry, “The Influence of Subjects Personality Traits on Personal Spatial Zones in a Human-Robot Interaction Experiment”, *Proc. IEEE Ro-man*, 2005.
- [10] D. Lambert, *Body Language*. HarperCollins, 2004.
- [11] L. R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, *Proc. of the IEEE*, vol. 77, no. 2, 1989.
- [12] S. Calinon and A. Billard, “Learning of Gestures by Imitation in a Humanoid Robot”, Dautenhahn, K. and Nehaniv, C.L. (eds.). *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions*. Cambridge University Press. 2005. In press.
- [13] T. Westeyn, H. Brashear, A. Atrash and T. Starner, “Georgia Tech Gesture Toolkit: Supporting Experiments in Gesture Recognition”, *Proc. 5th ICMI*, pp. 85-92, 2003.
- [14] J. M. Henderson and A. Hollingworth, “Global transsaccadic blindness during scene perception”, *Psychological Science*, vol.14, pp.493-497, 2003.

A Flexible Infrastructure for the Development of a Robot Companion with Extensible HRI-Capabilities*

J. Fritsch, M. Kleinehagenbrock, A. Haasch, S. Wrede, and G. Sagerer

Applied Computer Science, Faculty of Technology

Bielefeld University

33594 Bielefeld, Germany

jannik@TechFak.Uni-Bielefeld.DE

Abstract—The development of robot companions with natural human-robot interaction (HRI) capabilities is a challenging task as it requires incorporating various functionalities. Consequently, a flexible infrastructure for controlling module operation and data exchange between modules is proposed, taking into account insights from software system integration. This is achieved by combining a three-layer control architecture containing a flexible control component with a powerful communication framework. The use of XML throughout the whole infrastructure facilitates ongoing evolutionary development of the robot companion's capabilities.

Index Terms—robot architecture, human-robot interaction, system integration, distributed system, XML

I. INTRODUCTION

In recent years an increasing number of mobile robots have been constructed. They are spanning a wide range in terms of their interaction capabilities starting from tour guides (e.g., [1], [2]) across service robots which are used in office environments and for fetch-and-carry tasks (e.g., [3], [4]) up to personal robots with a stronger focus on natural interaction capabilities (e.g., [5], [6]).

Nearly all of the realized personal robots are research prototypes that integrate a variety of individual software components but reach up to now only a limited interaction quality. However, for a commercial success personal robots need to become true companions with more sophisticated human-robot interaction (HRI) capabilities that make the interaction as natural as possible. In addition, such a robot must also be capable of adapting itself to unknown environments and, therefore, it has to be able to acquire new knowledge in a lifelong learning process. Moreover, as humans are around, reactive control of the robot's hardware is important, too. Our goal is to build a robot that satisfies all these requirements so that it can be accepted by humans in their private homes: a so called *robot companion* [7].

For the realization of any complex robotic system a modular approach is essential [8]. For a robot companion a large number of software modules implementing the different aspects relevant for a smooth interaction behavior have to be developed and integrated. Because this is usually an iterative process in large-scale systems, it must be possible

to incorporate new modules and functionalities in a robot companion as they become available. Such a continuous extension of the robots functionality is usually very difficult as it not only requires modifying the control framework but also results in new data structures provided by the added components and possibly new control flows between modules. Therefore, the style of data flow in a modular system is also a crucial factor [9]. Another important aspect that is often neglected is the fact that the development of complex human-robot interaction architectures with many independent researchers involved is not only a matter of conceptual design but also a system engineering task.

We apply our mobile robot BIRON – the Bielefeld Robot Companion – to a specific application domain that currently gains increasing interest: the so-called *home tour scenario* (see also [7]). Here, a human introduces a newly bought robot all the objects and places in a private home relevant for later interaction between the human and the robot. For realizing such a system we have developed a generic robot control architecture to coordinate the activities of all integrated modules. The architecture allows a flexible extension of the overall system and is well suited to support natural human-robot interaction [10]. Our design decisions are motivated by earlier experiences gained from a previous implementation of the robot system [11] that lacked an advanced human-robot interface. The module communication on that robot had been realized using our former communication system DACS [12].

To provide a technical basis for building a robot companion, this paper focuses on the combination of our architectural methodology with the flexible, domain-independent, and easy-to-use XML-enabled Communication Framework (XCF) [13]. The resulting *System Infrastructure for Robot Companion Learning and Evolution* (SIRCLE) allows us to efficiently and transparently organize the realization and ongoing extension of our robot companion. We will show that through using XML as data format in the control component as well as for module communication the extension of the robot's software system with new modules can be realized very easily. This flexibility facilitates an agile software development process which has proven useful in software industry and results in a better ability to iteratively test intermediate versions of individual modules.

The paper is organized as follows: At first we discuss related work on robot architectures and their technical

*This work has been supported by the European Union within the 'Cognitive Robot Companion' (COGNIRON) project (FP6-IST-002020) and the 'VAMPIRE' project (IST-2001-34401) as well as by the German Research Foundation within the Graduate Programs 'Task Oriented Communication' and 'Strategies and Optimization of Behavior'.

realization in section II. Section III introduces our robot BIRON and in section IV we discuss requirements crucial for developing a robot companion. Subsequently, the XML-based communication framework is outlined in section V. Our robot control architecture and its implementation using XCF is described in section VI. Details of the control flow are presented in section VII and section VIII describes our experience with the proposed system infrastructure. The paper concludes with a short summary in section IX.

II. RELATED WORK

There are several mobile robot systems that integrate capabilities for human-robot interaction in their architecture. For example, Care-O-bot II [5] is a multi-functional robot assistant for housekeeping and home care. Its modular control architecture uses a central execution module to which all other modules are connected. Low-level modules are implemented in C++, but the task execution module is implemented in Python and makes use of the network communication supplied within Python. *HERMES* [4] is a humanoid service robot and can perform fetch-and-carry tasks. Although its system's core is behavior-based, the robot has a situation-oriented deliberative component to pursue long-term goals. It consists of proprietary hardware and runs a self written operating system that allows sending and receiving messages via different channels [14], while overall control is realized by using a finite state machine. *Jijo-2* [3] is intended to perform tasks in an office environment. Its system includes a reactive and an integrator layer, forming a hybrid architecture. The communication between the individual components is event-driven and based on low-level TCP/IP. *Lino* [6] serves as user interface to an intelligent home. To enable data exchange between all different components, a module-based software framework was developed, called the Dynamic Module Library. Here, each module has input and output ports that can be connected to each other to exchange data. On *ROBITA* [15], a robot that can participate in group conversations, all modules are connected by a priority-based coordination mechanism, which is based on a central blackboard [16]. A situation observation server monitors the blackboard and configures the overall system. This method allows to exchange and add new modules, but comes with the cost of the overhead for managing the blackboard.

Besides the architectures developed for the robots mentioned above there have been proposals for general robot architectures including application functionality. For example, *BERRA* [17], a three-layer architecture, is applied for robots performing fetch-and-carry and guiding tasks in the office domain. It is designed to provide scalability and a high degree of flexibility. Human-robot interaction covers mission acquisition only and, thus, user input is routed directly into a planner. The *LAAS* architecture [18] was originally designed for autonomous mobile robots, but is also used for other domains. It contains a central component, the so-called supervisor/executive, which coordinates data coming from the robot system and a planner, and commands from the operator. The focus is

on the execution of valid and safe commands which are verified in hard real-time. The Tripodal schematic control architecture [19] is also three-layered and enables robots to fulfill transportation tasks or work as a tour guide [20]. The overall system has a central process supervisor and its configuration is modeled by Petri nets which is advantageous for realizing coordination of parallel processes. However, temporal synchronization is not explicitly modeled and is realized by an ad hoc solution.

The robot architectures mentioned above are different from pure robot control architectures primarily providing interfaces to a robot's hardware like, e.g., *Player/Stage* [21] or Sony's *OPEN-R* [22]. Most of the architectures presented in this section can be classified as hybrid [23], often using a centralized component to control the individual system modules. However, these architectures support only limited interaction capabilities that are far from a natural interaction. Moreover, their extensibility does not allow to incorporate a large number of HRI components that will be needed for reaching a high interaction quality.

III. ROBOT HARDWARE

The software infrastructure proposed in this paper is running on our mobile robot BIRON (see Fig. 1). Its hardware platform is a Pioneer PeopleBot from ActivMedia with an on-board PC (Pentium III, 850 MHz) for controlling the motors and the on-board sensors as well as for sound processing. A second PC (Pentium III, 500 MHz) inside the robot is used for image processing. An additional laptop (Pentium M, 1.4 GHz) is used for speech processing and dialog control.

The two on-board PCs are linked via an 100 Mbit Ethernet LAN switch that is also equipped with an 11 Mbit wireless LAN. The laptop is linked to this switch wirelessly, but can also be mounted on the robot for full autonomy. All three computers are running Linux.

A pan-tilt color camera (Sony EVI-D31) is mounted on top of the robot at a height of 141 cm for acquiring images of the upper body part of humans interacting with the robot. Two AKG far-field microphones which are usually used for hands free telephony are located at the front of the upper platform at a height of 106 cm, right below the touch screen display. The distance between the microphones is 28.1 cm. A SICK laser range finder is mounted at the front at a height of 30 cm. As additional interactive device a 12" touch screen display is provided on the robot.



Fig. 1. BIRON.

IV. REQUIREMENTS FOR DEVELOPING A COMPANION

To support progressive development of a robot companion's HRI capabilities, the functional requirements *mod-*

ularity, communication, module coordination, as well as knowledge representation and acquisition motivated in the introduction must be supported by the proposed system infrastructure. Additionally, several non-functional requirements play an important role that are discussed in the following.

One fundamental task for communication frameworks in the robotics domain is the ability to distribute modules across different processing nodes in order to guarantee fast system responses [24], [25]. This applies especially to large-scale systems like robot companions. However, most robotic researchers are no middleware experts, prohibiting the native use of, e.g., CORBA-based solutions. Communication frameworks built on top of CORBA [24] try to encapsulate complexity with domain-specific class libraries, which often complicates their use in system architectures of other domains. Thus, one requirement for a generic communication framework is the ability to enable researchers to easily build a distributed robot architecture. Additionally, it should allow for frequent integration cycles. To achieve these criteria, *simplicity, usability* and *standards compliance* are essential.

Furthermore, specifications change frequently in a research prototype. Thus an important feature is the *flexibility* of the communication framework. The impact of interface changes on an existing system architecture should be minimal to avoid the versioning problem [26]. Another important requirement that benefits directly from high usability and flexibility is *rapid prototyping*. Consequently, iterative development should not only be supported for single modules but also for the integrated system. Wrong directions in system evolution can more easily be identified if integration is performed on a regular basis starting at an early stage. For large-scale systems, software engineering research has shown that decoupling of modules is very important. Thus, the framework should support *low coupling* of modules. This facilitates not only independent operation of components but also minimal impact of local changes on the whole system. With a framework that adheres to low coupling, *debugging and evaluation* of a running system architecture can be supported more easily.

V. XML ENABLED COMMUNICATION FRAMEWORK

Taking into account the above mentioned non-functional as well as the basic functional requirements like module communication with flexible knowledge representations and data management, we developed XCF [27] in the context of cognitive vision and robotics. In the following we will present the fundamental concepts and show how these help to build an extensible robot architecture infrastructure.

Since it is very flexible, easy to learn and suited for abstract concept descriptions, XML was chosen as a basis to describe content transmitted, stored, and processed by the various robot modules. Note, that XML is not only used as a data exchange protocol as in XML-RPC-based solutions [25] that usually produce a lot of overhead through text-based representation of binary data and parameter encoding rules. Instead, we developed XML vocabularies

for symbolic robotic data (e.g., states, events, objects, etc.). The instance documents then contain directly the semantical information that is accessed and selected using the standardized XQuery/XPath mechanisms.

Meta-information, e.g., about data types, is kept separate in the corresponding XML schema files and is not encoded in the instance document itself. Data type specification with XML schemas has several advantages in comparison to traditional programming language constructs. First of all, the data types are independent from specific programming languages. Even so, tools for using them are available for almost every platform. Furthermore, XML schemas are able to specify content models and ranges of allowed values in great detail. Providing fine grained sets of semantically grouped declarations in separate schemas with associated XML namespaces makes them reusable throughout different systems. Complex schemas for individual modules can then easily be composed out of these basic type libraries, only adding specific complex types. If taken into account, extensibility of data types is possible with schema evolution. Even complex grammars for components capable of interpreting and validating XML documents originating from different robot modules (e.g., the execution supervisor presented in section VII is an example for such a module) are easy to compose and well understandable with a sophisticated schema hierarchy.

Using XML for knowledge representation and schemas for data type specification, the Internet Communication Engine (ICE) [28] was chosen as technical basis of our framework. ICE offers similar functionality as CORBA, but with a much more lightweight approach. The ICE core manages communication tasks using an efficient protocol, provides a powerful thread mechanism and additional functionality that supports scalability. Similar to CORBA, ICE also relies on pre-compiled proxy objects. Unlike CORBA, ICE has no explicit dynamic invocation interface by itself.

On top of the ICE library XCF was developed to provide an easy to use middleware for building distributed object-oriented architectures that can efficiently exchange XML and referenced binary data (e.g., images). The referenced binary data structures are transmitted natively in a composite object together with the corresponding XML message. This combines the flexibility of XML with the efficiency of low-level communication semantics for large amounts of binary data. The XCF core itself features a pattern-based design and offers communication semantics like publisher/subscriber and (a)synchronous remote procedure calls/method invocations as well as event channels. All XCF objects and exposed methods can be dynamically registered at runtime.

Since data types are specified using XML schema as explained above, runtime type safety can be ensured. System introspection is directly supported through the implementation of the interceptor pattern [29] that helps in debugging and monitoring a running distributed system.

XCF conforms at least to the following transparency levels which are important for communication frame-

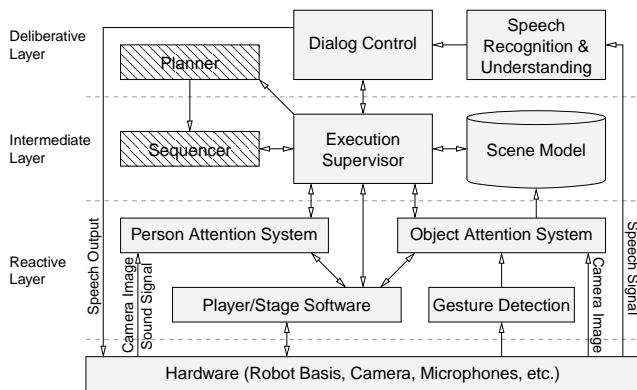


Fig. 2. Interaction architecture of BIRON.

works [30]: *Access transparency* is provided by the XCF core, where the implemented dispatcher service realizes *location transparency*. *Concurrent access* of multiple clients on one server is transparently handled by a specific worker thread pool. Additionally, the use of monitor threads provides *migration* and *error transparency* for computational modules.

To address the issue of data management, the active memory concept and implementation [31] is applied for use in our robotic framework. On top of a native XML database library [32], a server architecture was implemented that allows processing of the above mentioned data messages consisting of XML and referenced binary data. Thus, not only XML but also binary data can be shared by several robotic modules in parallel. For both kinds of data, powerful standard DBMS methods like insert, update, remove and query are exposed. Node selection and referral is based on XPath statements. Closely coupled to the knowledge representation a subscription model for distributed event listeners was realized, so that memory events can trigger registered processes that are interested in specific data and/or memory actions.

Though XCF is realized in C++ for performance reasons, there are also Matlab and Java bindings for rapid prototyping. It provides an easy to use basis for distributed processing in an asynchronous, decoupled fashion. Declarative, name-based selection of XML nodes with XPath expressions helps in building data types that can be extensible and in building systems that will not break if modifications occur. The following sections show how these foundations are used to realize a flexible and extensible architecture for human-robot interaction.

VI. ARCHITECTURE

In principle our system is based on a three-layer architecture [33], as it is the most flexible way to organize a system which integrates autonomous control and human-robot interaction capabilities [10]. An overview of the resulting architecture can be seen in Fig. 2.

The most important component concerning the structure of the proposed architecture is a central *execution supervisor* (see section VII). The functionality of this execution

supervisor is similar to the so-called *sequencer* used for ATLANTIS [34] where it coordinates the operations of the modules responsible for deliberative computations rather than vice versa. This is contrary to most hybrid architectures where a deliberator continuously generates plans and the reactive plan execution mechanism just has to make sure that a plan is executed until a new plan is received.

The main modules in the deliberative layer are the planner and the dialog control. The planner is responsible for generating plans for, e.g., navigational tasks. The dialog control carries out dialogs to obtain instructions given by a human interaction partner via the speech understanding system [35]. It is also able to manage interaction problems and to resolve ambiguities by consulting the user. The dialog control sends valid instructions to the execution supervisor which is located in the intermediate layer of our architecture. Because the dialog control is directly connected to the central component, ambiguities that might arise from modules in the reactive layer can also be resolved by dialog. For this purpose corresponding enquiries from the reactive layer are routed through the execution supervisor. Thus, results from HRI are made available centrally in the architecture instead of routing them only to a planner for pure mission acquirement like it is done in many other robot architectures (see, e.g., [17], [19]).

The sequencer also resides in the intermediate layer of our architecture. It is responsible for decomposing plans provided by the planner, as the execution supervisor can only handle single commands. A scene model for maintaining knowledge representations completes this layer.

The person attention system [36] represents the robots main reactive feedback control mechanism and is therefore located in the reactive layer. It detects potential communication partners among persons present in the vicinity of the robot. It is configured by the execution supervisor to show different behaviors, e.g., to look at all people in its surrounding or to track a specific communication partner. However, the person attention system does not directly control the robot's hardware. This is done by the *Player/Stage* software [21] which provides a clean and simple interface to the robot's sensors and actuators. Even though we currently use this software to control the hardware directly, the controller can easily be replaced by a more complex one which may be behavior-based to also include, e.g., obstacle avoidance.

Besides the person attention system an object attention system is located in the reactive layer. The execution supervisor can shift control of the robot from the person attention system to the object attention system in order to focus objects referred to by the user. The object attention is supported by a gesture detection module which recognizes deictic gestures [37]. Combining spoken instructions and a deictic gesture allows the object attention system to acquire visual information of a referenced object. This information is sent to the scene model in the intermediate layer.

The scene model stores information about objects introduced to the robot for later interactions. This information

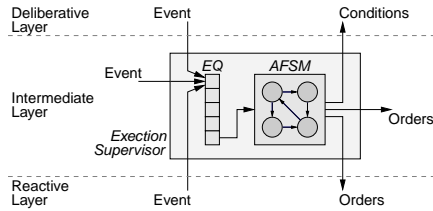


Fig. 3. Schema of the execution supervisor. It contains an event queue (EQ) and is controlled by an augmented finite state machine (AFSM).

includes attributes like position, size, and visual information of objects provided by the object attention module. Besides, additional information given by the user is stored in the scene model, e.g., a phrase like “This is my coffee cup” indicates owner and use of an object to learn.

In order to satisfy certain system safety requirements, modules should fail perceptibly as in a real world robot failures cannot be excluded. We realized this feature by messages which are initiated by the main loop of each module and sent in fixed intervals to modules being in communication with this module. If a module does not receive these messages anymore, it can determine that the corresponding sender stopped working correctly. In this case corrective actions can be taken to recover from the failure. If recovery is not possible then the robot is at least able to ask the user to call technical support.

VII. EXECUTION SUPERVISOR

The execution supervisor is the central part of our architecture and is designed to be as generic as possible. In order to achieve this requirement the execution supervisor interprets no data at all: It either configures modules of the system at runtime-based on received events containing needed parameters, or it routes specific data to modules which are responsible for processing the data. Since all information is carried by XML documents, the execution supervisor even does not need to distinguish between the data structures it receives.

The execution supervisor receives data in the form of events from all modules connected to it. Because these events occur asynchronously and latencies in the communication have to be considered, the execution supervisor uses an *event queue* (see Fig. 3). All incoming events contain timestamps indicating when they were created. They are inserted in the queue ordered by their timestamps. The events are handled in turn, starting with the oldest one.

The event queue is also used to synchronize events. For example, a person can only become the robot’s current communication partner if the person attention system signals that it has detected a potential communication partner and the dialog control notifies the execution supervisor that it has received a corresponding speech input. Only if both events arrive at the execution supervisor in a certain interval of time, it is assumed that these events belong together. This is verified by a lifetime entry in the corresponding events which describes how long an event remains valid. For a more detailed explanation please see [10].

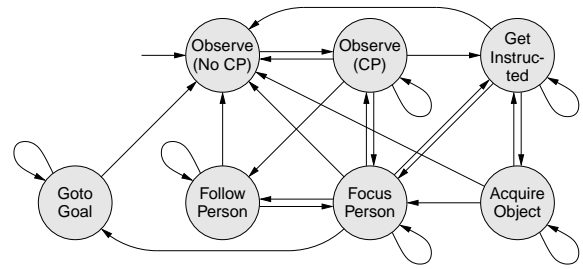


Fig. 4. Augmented finite state machine of the execution supervisor.

In order to process the events stored in the event queue, the execution supervisor is controlled by an *augmented finite state machine (AFSM)*. Each event corresponds to a transition in the AFSM. Thus, transitions from one state to another can only be triggered by events. When a transition is executed, the corresponding event is deleted from the event queue. As the execution supervisor is not intended to interpret any data, the AFSM is not very complex (see Fig. 4). Therefore, the structure of the AFSM is also specified in XML. As a consequence, no special-purpose language like, e.g., RAPS [38] is necessary. Defining the AFSM in XML results in a clear representation which allows to quickly restructure or extend the execution supervisor without recompiling it. This concept also allows us to modify the execution supervisor at runtime. As new XML documents can be handled directly by default their integration is straightforward.

The finite state machine is augmented in so far, that with every transition which is executed, a specific *action* is performed. These actions are for configuring the system by emitting two specific types of messages: *conditions* and *orders*. These messages include parameters needed by the addressed receivers. The parameters are supplied by the event which initiated the corresponding action. Orders are sent to all modules in the intermediate and reactive layer in order to reconfigure the system. Conditions are sent to modules in the deliberative layer which inform these modules about the internal state of the overall system. This form of communication reflects the hierarchical structure of the architecture: Orders are sent ‘downwards’, while conditions are sent ‘upwards’.

An example on how the execution supervisor routes information from one module to another one can be seen in Fig. 5. Here, the execution supervisor receives an event from the person attention system, which contains new data about a communication partner. Depending on the overall system state, this event initiates a specific condition, which is then sent to the dialog control module. A copy of the data from the event is included in this condition.

Altogether, the execution supervisor processes events from different modules and also ensures that certain events have to arrive in a certain time interval before the contained AFSM changes to a specific state. This reflects that the execution supervisor is in control of the overall system. The dialog control agent can give advices, but if the components of the remaining layers do not provide corresponding

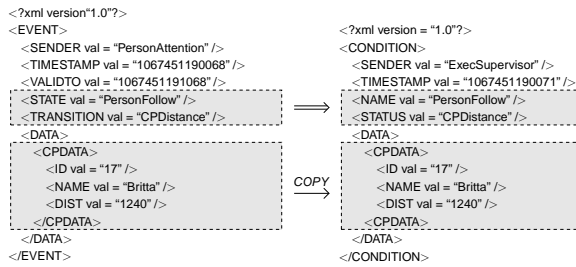


Fig. 5. Routing example: Data of an event from the person attention system are propagated to the dialog control via a specific condition.

information advices are rejected. For example, the robot may perceive instructions from a radio, but it will not start an interaction as no person can be detected. This makes the interaction with the overall system more robust.

Due to the low complexity, our concept of the execution supervisor also scales up if the system is extended. Generally, only one more state is needed when a new functionality or module is added.

VIII. EXPERIENCE

The presented system infrastructure SIRCLE was successfully applied to our mobile robot BIRON. With several researchers contributing to this system, the proposed solution proved its suitability for the realization of the complex robot architecture presented in section VI.

The module integration task directly benefits from the fact that XCF is very easy to use and the applied concepts are highly standards-based. Furthermore, a central system view utilizing the active system introspection supports debugging and evaluation of the running system.

Loose coupling of modules and the declarative style of accessing system data paid off in ease of modification and the ability to let the system architecture evolve over time as we integrated new modules and data types. For example, a first prototype of the object attention system was added just recently and required on the architectural level only a modification of the execution supervisor's configuration file and thus demonstrating iterative development. Additionally, loose coupling of modules as well as error and migration transparency yield increased robustness of the overall system.

With respect to the performance of the overall system, the distribution and coordination of modules using SIRCLE has resulted in a low response time allowing for more natural human-robot interaction. In order to obtain quantitative results of the overall system behavior corresponding time measurements were carried out using a typical interaction example. The results of this evaluation are presented in Fig. 6 in a sequence diagram. All interactions with BIRON follow this pattern in principle, even though the duration for the speech processing might vary. The main reason for this is that the time needed to process speech input depends on the length of the utterance given by the user.

In Fig. 6 it can be seen how much time is needed for the most important processing phases and that the speech

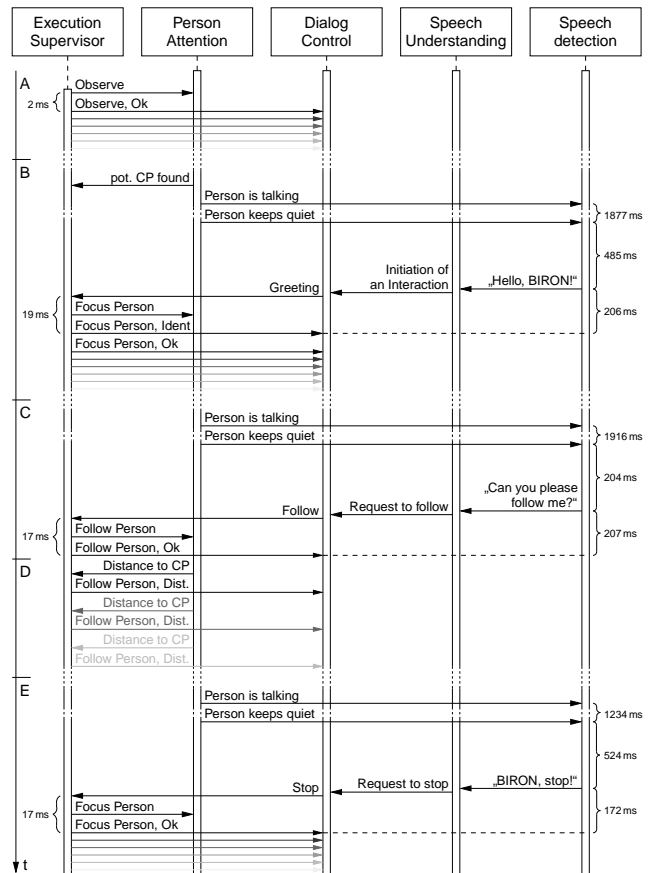


Fig. 6. Qualitative Representation of the time course of the module communication on BIRON considering an interaction example. Arrows getting brighter indicate messages that are sent continuously. Different phases: A: The system is started. B: A person approaches BIRON and greets the robot. C: The person asks BIRON to follow her. D: BIRON follows the person. E: The person asks BIRON to stop following.

detection takes most of the processing time. Subsequently the system usually needs only around 200 ms until it produces an answer to the user. According to this duration, the execution supervisor takes only a very small amount of time to reconfigure the system: less than 20 ms. Time needed for module communications is also very small and shows that the advantage of increased processing power in a distributed system exceeds the cost for module communication. Transferring a message from one to another module takes only a few milliseconds in average, but depends on whether it is routed via WLAN or not. A more fine grained examination is difficult as time measurements slow down the system and falsify results. However, the results show that the software infrastructure on BIRON enables efficient module interaction that leads to a reactive overall system.

The suitability of our system infrastructure SIRCLE was proven by presenting two BIRON robots at the exhibition of the Information Society Technologies Event (IST Event) in November 2004 in The Hague, The Netherlands [39]. The robots were presented to the public at all three exhibition days, twice for 9 hours and once for 6 hours. They were instructed to either follow the user or to look at some objects. The robots worked continuously and robust.

Altogether visitors watched the presentation of the BIRON systems with great interest and the performance of the robots was consistently denoted as impressing.

IX. SUMMARY

In this paper we presented SIRCLE, a system infrastructure providing a software platform for a robot companion exhibiting powerful capabilities in human-robot interaction. Our approach combines the XML-enabled Communication Framework presented in section V with our three-layer architecture for controlling module operation and communication as shown in sections VI and VII. The overall concept proved its suitability during the ongoing work on our robot companion BIRON. Targeting not only the functional requirements, our infrastructure features simplicity, standards compliance, and extensibility which results in an agile development process and, ultimately, in a robot companion with more natural human-robot interaction capabilities. The suitability of our approach was proven by a quantitative system evaluation and the successful presentation of our two BIRON robots at the IST Event 2004.

REFERENCES

- [1] S. Thrun, *et al.*, *Probabilistic Algorithms and the Interactive Museum Tour-Guide Robot Minerva*, *International Journal of Robotics Research*, vol. 19, no. 11, pp. 972–999, 2000.
- [2] N. Tomatis, *et al.*, “Building a Fully Autonomous Tour Guide Robot: Where Academic Research Meets Industry,” in *Proc. Int. Symp. on Robotics*, Stockholm, Sweden, 2002.
- [3] H. Asoh, *et al.*, *Jijo-2: An Office Robot that Communicates and Learns*, *IEEE Intelligent Systems*, vol. 16, no. 5, pp. 46–55, 2001.
- [4] R. Bischoff and V. Graefe, “Demonstrating the Humanoid Robot HERMES at an Exhibition: A Long-Term Dependability Test,” in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems; Workshop on Robots at Exhibitions*, Lausanne, Switzerland, 2002.
- [5] B. Graf, M. Hans, and R. D. Schraft, *Care-O-bot II—Development of a Next Generation Robotic Home Assistant*, *Autonomous Robots*, vol. 16, no. 2, pp. 193–205, 2004.
- [6] B. J. A. Kröse, *et al.*, “Lino, the User-Interface Robot,” in *European Symposium on Ambient Intelligence (EUSAI)*, Veldhoven, Netherlands, 2003, pp. 264–274.
- [7] “The Cognitive Robot Companion,” 2004, <http://www.cogniron.org>.
- [8] J. K. Rosenblatt, “DAMN: A Distributed Architecture for Mobile Navigation,” in *Proc. AAAI Spring Symp. on Lessons Learned from Implemented Software Architectures for Physical Agents*. Stanford, CA: AAAI/MIT Press, 1995, pp. 167–178.
- [9] È. Coste-Manière and R. G. Simmons, “Architecture, the Backbone of Robotic Systems,” in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 1, San Francisco, CA, 2000, pp. 67–72.
- [10] M. Kleinhagenbrock, J. Fritsch, and G. Sagerer, “Supporting Advanced Interaction Capabilities on a Mobile Robot with a Flexible Control System,” in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, vol. 3, Sendai, Japan, 2004, pp. 3649–3655.
- [11] J. Fritsch, *et al.*, *Multi-Modal Anchoring for Human-Robot-Interaction, Robotics and Autonomous Systems; Special issue on Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems*, vol. 43, no. 2–3, pp. 133–147, 2003.
- [12] G. A. Fink, N. Jungclaus, F. Kummert, H. Ritter, and G. Sagerer, “A Distributed System for Integrated Speech and Image Understanding,” in *Int. Symp. on Artificial Intelligence*, Cancun, Mexico, 1996, pp. 117–126.
- [13] S. Wrede, J. Fritsch, C. Bauckhage, and G. Sagerer, “An XML Based Framework for Cognitive Vision Architectures,” in *Proc. Int. Conf. on Pattern Recognition*, vol. 1, Cambridge, UK, 2004, pp. 757–760.
- [14] R. Bischoff, “Recent Advances in the Development of the Humanoid Service Robot HERMES,” in *Proc. European Workshop and Masterclass on Advanced Robotics Systems Development (EUREL)*, vol. 1, Manchester, UK, 2000, pp. 125–134.
- [15] K. Kim, Y. Matsusaka, and T. Kobayashi, “Inter-Module Cooperation Architecture for Interactive Robot,” in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, vol. 3, Lausanne, Switzerland, 2002, pp. 2286–2291.
- [16] B. Hayes-Roth, *A Blackboard Architecture for Control, Artificial Intelligence*, vol. 26, no. 3, pp. 251–321, 1985.
- [17] M. Lindström, A. Orebäck, and H. I. Christensen, “BERRA: A Research Architecture for Service Robots,” in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 4, San Francisco, CA, 2000, pp. 3278–3283.
- [18] F. F. Ingrand and F. Py, “An Execution Control System for Autonomous Robots,” in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 2, Washington DC, USA, 2002, pp. 1333–1338.
- [19] G. Kim, W. Chung, M. Kim, and C. Lee, “Tripodal Schematic Design of the Control Architecture for the Service Robot PSR,” in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 2, Taipei, Taiwan, 2003, pp. 2792–2797.
- [20] G. Kim, W. Chung, M. Kim, and C. Lee, “Implementation of Multi-Functional Service Robots Using Tripodal Schematic Control Architecture,” in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 4, New Orleans, LA, 2004, pp. 4005–4010.
- [21] B. P. Gerkey, R. T. Vaughan, and A. Howard, “The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems,” in *Proc. Int. Conf. on Advanced Robotics*, Coimbra, Portugal, 2003, pp. 317–323.
- [22] “Sony’s OPEN-R SDK,” <http://openr.aibo.com>.
- [23] M. J. Mataric, *Learning in Behavior-Based Multi-Robot Systems: Policies, Models, and Other Agents*, *Cognitive Systems Research; Special issue on Multi-disciplinary studies of multi-agent learning*, vol. 2, no. 1, pp. 81–93, 2001.
- [24] S. Knoop, S. Vacek, R. Zöllner, C. Au, and R. Dillmann, “A CORBA-Based Distributed Software Architecture for Control of Service Robots,” in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, vol. 3, Sendai, Japan, 2004, pp. 3656–3661.
- [25] P. Kiatisevi, V. Ampornaramveth, and H. Ueno, “A Distributed Architecture for Knowledge-Based Interactive Robots,” in *Proc. Int. Conf. on Information Technology for Application (ICITA)*, Harbin, China, 2004, pp. 256–261.
- [26] D. C. Schmidt and S. Vinoski, “Object Interconnections: CORBA and XML, Part 1: Versioning,” 2003, <http://www.cs.wustl.edu/~schmidt/report-doc.html>.
- [27] “XCF at Sourceforge.Net,” 2004, <http://xcf.sourceforge.net>.
- [28] “The Internet Communications Engine, ZeroC Inc.” 2003, <http://www.zeroc.com/ice.html>.
- [29] D. C. Schmidt, M. Stal, H. Rohnert, and F. Buschmann, *Pattern-Oriented Software Architecture*. John Wiley & Sons Ltd., 2000, vol. 2: Patterns for Concurrent and Networked Objects.
- [30] A. S. Tanenbaum and M. van Steen, *Distributed Systems: Principles and Paradigms*, 1st ed. Prentice Hall, 2002.
- [31] S. Wrede, M. Hanheide, C. Bauckhage, and G. Sagerer, “An Active Memory as a Model for Information Fusion,” in *Proc. Int. Conf. on Information Fusion*, vol. 1, Stockholm, Sweden, 2004, pp. 198–205.
- [32] “Berkely DB XML, Sleepycat Software.” 2004, <http://www.sleepycat.com/products/xml.shtml>.
- [33] E. Gat, “On Three-Layer Architectures,” in *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems*. Cambridge, MA: MIT Press, 1998, ch. 8, pp. 195–210.
- [34] E. Gat, “Integrating Planning and Reacting in a Heterogeneous Asynchronous Architecture for Controlling Real-World Mobile Robots,” in *Proc. Nat. Conf. on Artificial Intelligence (AAAI)*. San Jose, CA: AAAI/MIT Press, 1992, pp. 809–815.
- [35] A. Haasch, *et al.*, “BIRON – The Bielefeld Robot Companion,” in *Proc. Int. Workshop on Advances in Service Robotics*. Stuttgart, Germany: Fraunhofer IRB Verlag, 2004, pp. 27–32.
- [36] S. Lang, *et al.*, “Providing the Basis for Human-Robot-Interaction: A Multi-Modal Attention System for a Mobile Robot,” in *Proc. Int. Conf. on Multimodal Interfaces*. Vancouver, Canada: ACM, 2003, pp. 28–35.
- [37] N. Hofemann, J. Fritsch, and G. Sagerer, “Recognition of Deictic Gestures with Context,” in *Pattern Recognition; 26th DAGM Symposium, Tübingen, Germany, Aug./Sep. 2004. Proceedings*. Heidelberg, Germany: Springer-Verlag, 2004, vol. 3175, pp. 334–341.
- [38] R. J. Firby, “Adaptive Execution in Dynamic Domains,” Ph.D. dissertation, Yale University, New Haven, CT, 1989.
- [39] “Information Society Technologies Event (IST Event),” Nov. 2004, http://europa.eu.int/information_society/istevent/2004/index_en.htm.

Methodological Issues Using a Comfort Level Device in Human-Robot Interactions

Kheng Lee Koay, Michael L. Walters, Kerstin Dautenhahn
Adaptive Systems Research Group, School of Computer Science
University of Hertfordshire
College Lane, Hatfield, Herts, AL10 9AB, United Kingdom
{K.L.Koay,M.L.Walters,K.Dautenhahn}@herts.ac.uk

Abstract – This paper introduces a handheld Comfort Level Device to measure subjects’ comfort levels in human-robot interaction experiments. We discuss methodological issues of using the device in an exploratory HRI study where subjects were asked to use the device to indicate their subjective comfort level throughout the experiment. The recorded comfort data were time stamped for synchronization and analysis purposes in conjunction with the video footage to help identify certain situations in the HRI trials where subjects felt uncomfortable. In order to provide a proof-of-concept for the suitability of the handheld Comfort Level Device for HRI studies we analyzed the data for seven selected subjects. These examples show that our method helped identifying robot behaviors that subjects felt uncomfortable with. We demonstrate that the device revealed certain uncomfortable states that are visually hidden. Limitations of the device and possible implications for future work conclude the paper.

Index Terms – Human-Robot Interaction, Social Robot, Social Interaction, Comfort Level Device.

I. INTRODUCTION

In human-inhabited social environments, behaviours or tasks that robots exhibit or perform will result in certain behaviours or responses of humans. Therefore it is essential to understand the relationship between human and robot behaviours in order to create *social robots that humans feel comfortable with*. The issue of human social acceptance has led to studies that concentrate on the human-centered perspective, where it is essential to include the human in the loop in order to understand what attributes (i.e. behaviour styles, appearance etc.) of robots elicit interactions that are comfortable from the perspective of humans [1,2,3,4]. The research reported in this paper is part of the European project COGNIRON and studies robot companions in a home setting. While such a robot needs to perform and provide assistance for certain useful tasks [5], it should also behave in a socially acceptable manner.

Two main strategies are commonly used for evaluating human-robot interaction from a human subjects’ perspective: 1) questionnaires, e.g. used in [5], and 2) analysis of video footage recording the interactions, e.g. [6,7,8]. The latter is more appropriate for scenarios where e.g. verbal inquiry may

be impossible (e.g. in the case of non-verbal subjects) [7,8], too intrusive, or might strongly bias the results [9].

For video analysis in our study, a video annotation tool was used to annotate and catalogue specific behaviours of interest from the video footage. Drawback of video analysis is that it is a very time consuming method and that it requires inter-rater reliability tests. Trained video observers are necessary to perform the video analysis. However, there is no guarantee that they will be able to observe all relevant behaviours, let alone subjects’ comfort levels which might, if at all, be revealed through language or subtle cues (e.g. facial expressions or utterances indicating discomfort or comfort). Therefore, ‘feeling comfortable or uncomfortable’ is not necessarily expressed clearly enough that it can be detected from observing video footage. Individual differences in subjects’ *expressiveness*, as well as the problem of being able to monitor the subject’s face, body movements and utterances continuously during the experiments have encouraged us to pursue an alternative.

In human-computer interaction and robotics, biofeedback sensors measuring physiological variables such as heart beat or skin conductance etc. have been investigated¹. However, the signal processing required for detecting affect and other internal states is often extensive and sensors need to be attached to the subject. Deriving a high-level concept such as ‘comfort’ from rich physiological data is not straightforward, although subjects are very familiar with assessing their own subjective ‘comfort level’. Thus, we decided to try to directly measure a subject’s comfort level via a simple device where subjects use a continuous scale to judge their current comfort level throughout an HRI interaction trial. This led us to the posing of two research questions addressed in the present paper:

RQ1: *Can a simple handheld device be used as a tool for helping researchers identify subjects’ comfort level?*

RQ2: *Can a visually hidden uncomfortable state be identified through the use of the Comfort Level Device?*

¹ E.g., D. Kulic and E. Croft, “Estimating intent for human-robot interaction”, *Proc. IEEE Int. Conf. on Advanced Robotics*, 2003, pp. 810-815; R. Rani, N. Sarkar, C. Smith, L. Kirby, “Anxiety detecting robotic systems – Towards implicit human-robot collaboration”, *Robotica*, 22(1): 85-95, 2004.

II. HUMAN-ROBOT INTERACTION TRIALS

The exploratory study involved single human subjects in a simulated living room scenario. It was carried out at the University of Hertfordshire premises between July and August 2004. This study was conducted using a commercially available, human-scaled, PeopleBot™ robot. The main aim of the study was to evaluate, in a task oriented living room scenario, different social behaviour and interaction styles of the PeopleBot™ robot from a human-centred perspective. A sample of 28 adult volunteers was recruited from the University of Hertfordshire, balanced for gender, background, and familiarity with technology. All subjects completed consent forms and were not paid for participation.

A. Experimental Design

Experimental Setup - The Simulated Living Room

The original room measured 8.5 x 4.75m and was partitioned off at one end to form an area that served as a control area for the Wizard-of-Oz [10,11] operators and provided space for the control, network and recording equipment. The room was decorated as a simulated living room.

B. The Experimental Procedure

The experiment was supervised by an experimenter who introduced and explained the trials to the subject. Each single subject spent about 50 minutes in the simulated living room with only the robot and the experimenter present who interfered as little as possible with the robot trails. The following phases of the experimental procedure are relevant to the present paper.

Introduction: A general welcome phase where the robot was introduced to the subject when they entered the simulated living room. An information sheet was given to the subject to read along with a consent form to be signed, then questionnaires were completed. The robot moved around the room whilst the subject completed these initial questionnaires in order to familiarize the subject with the robot.

Comfort Level Device: Before subjects proceeded to the main trial, they were given a Comfort Level Device (Fig. 1) and were asked to try it out and operate it a few times (for calibration purposes and in order to provide an opportunity for the subject to get accustomed to the device²). Next, they were told to use it throughout the main trial to indicate their comfort level during the trial (see section III). A subset of the data collected in this way during the trials formed the basis of this paper³.

² The handheld device might provide an additional potential source of discomfort. We tried to reduce this effect by allowing time for the subject to get used to device. Any potential additional discomfort is likely to be present during the whole trial, and thus less likely to influence the *changes in the levels of comfort/discomfort* which were our primary concern. Focussing on changes in the comfort levels has a second advantage: it makes the data more independent of any 'moods' that a particular subject might be in e.g. on a particular day, assuming that such moods are persistent over a longer period of time. However, these issues merit further investigation.

³ In terms of the experimental design of our study, we would like to make the following remarks: It would indeed be interesting to see how subjects in a control group, not using the handheld device, would behave. However, the

Main Trial: The main trial consisted of two tasks, a Negotiated Space Task and an Assistance Task. The Negotiated Space Task involved the robot moving in the room while the subject went through a pile of books placed on the table, remembering one title at a time, walking over and writing down each title on the whiteboard. The Assistance task involved the subject sitting at the table, copying the book titles from the whiteboard onto a piece of paper and underlining specific letters with a red/highlighter pen. The robot was responsible for bringing the missing red/highlighter pen to the table. The two tasks were chosen as they match two key scenarios studied in the COGNIRON project [12]. At the end of these two task scenarios, the subject completed a robot personality questionnaire. The Main trial was then repeated.

Final Phase: The final phase involved the subjects completing several questionnaires.

III. RESULTS FROM COMFORT LEVEL DEVICE

We built a handheld comfort level monitoring device that would allow subjects to indicate their internal comfort level during the experiment (Fig. 1).



Fig. 1 Photograph of the handheld Comfort Level Device.

The device uses a slider control, located at one edge of the box, to receive users' comfort level feedback. The slider can be moved easily by the subjects using either a thumb or finger to indicate their comfort level. The slider scale was marked on one end of the slider with a happy face, to indicate the subject was comfortable with the robot's behaviour, and a sad face on the other end, to indicate discomfort with the robot's behaviour. The device used a 2.4GHz radio signal data link to

primary purpose of our study was to identify whether the handheld device could be used to relate subjects' subjective judgements of comfort/discomfort with observable behaviour. A group of subjects using other, more sophisticated and expensive (e.g. physiological) devices to identify discomfort could serve as a suitable control group. However, those alternative devices were not available to us, and, it is not clear how to easily deduce comfort/discomfort from physiological data. Asking for vocalisations (e.g. "I don't feel comfortable now", or verbal ratings on a scale from one to ten) did not seem appropriate either since it would have interfered with the reading/writing tasks that the subjects were performing. Also, moving a slider with one finger seemed easier to us compared to the effort required in order to pinpoint verbally exact moments of discomfort. Vocalizations would also not be able to provide fine graded quantitative data. Note, our primary aim is to develop a reliable Comfort Level Device for human-robot trials. Thus, a control group involving human-human interaction, instead of human-robot interaction, did not seem suitable either. Our main motivation was to use a simple, very inexpensive device, that can easily be replicated by any talented person with certain engineering skills, and to propose a simple data analysis technique respectively.

send numbers representing the slider position to a PC mounted receiver, which recorded the slider position approximately 10 times per second. The data was time stamped and saved in a file for later synchronisation and analysis in conjunction with the video material. The data downloaded from the handheld subject Comfort Level Device was saved and plotted on a series of charts. However, unexpectedly, the raw data was heavily corrupted by static from the network cameras used to make video recordings of the session (see Fig. 2). We thus developed a method that can digitally clean up this static noise, explained in the next section.

A. Noise Filtering

In this section we describe a simple technique for noise reduction in the data⁴. By carefully analysing the raw comfort data, plotted against time (e.g. Fig. 2), we found that it was difficult to distinguish the static noise from the actual comfort data at certain regions of a plot (e.g. the region at time 14:37:41). To overcome this problem, we decided to spread the data points out by plotting the raw comfort level data along the x-axis that was incremented by one data point per step (see Fig. 3). We performed the same plotting method for the subjects' calibration data (see Fig. 4). Next, by comparing the raw comfort level data with the subject's calibration data, we noticed that the characteristics of the static noise were very different from a natural human sliding movement shown in Fig. 4. The raw comfort data contained a lot of random spikes (which were characteristic of static noise) in addition to what appeared to be the subject's actual comfort level profile.

To filter out these random spikes, we decided to use the user calibration data as a reference to determine a threshold value that can be used in our filtering process to prune these random spikes from the raw comfort data. The threshold value was determined by searching through the calibration data to obtain the maximum difference between two data points. The idea was to use the maximum difference between two data points as a threshold value that represents the actual maximum linear velocity the subjects moved the slider under normal conditions. We assumed that only static noise can cause a difference between two points in the raw data exceeding the threshold value.

By using the threshold value, we then scanned through the raw data and replaced the static noise (e.g. p_i) with their previous non-static noise data point (e.g. p_{i-1}). Note that the threshold value varies with subjects; therefore it was essential to determine each subject's threshold value separately through their calibration data during the filtering process.

Figure 5 illustrates the actual comfort data profile after the filtering process of the raw data (Fig. 2) using a threshold value of 51.

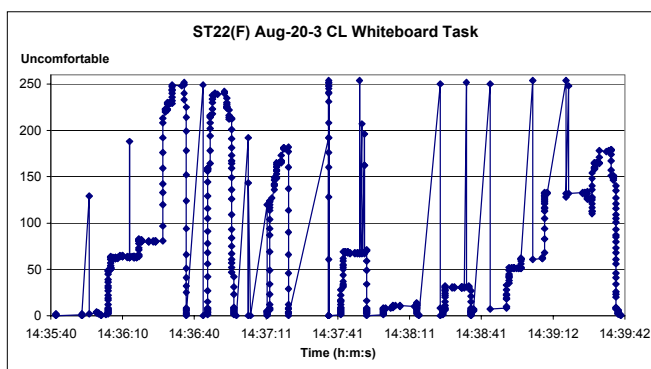


Fig. 2 Raw comfort level data plotted against time.

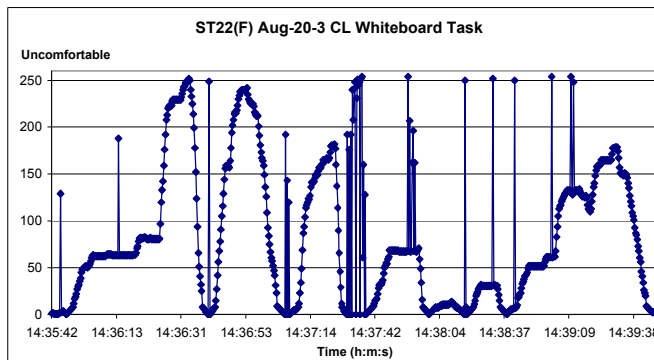


Fig. 3 Raw comfort level data plotted on the x-axis which increments by one data point per step. The time is stamped on the graph every 27 data points.

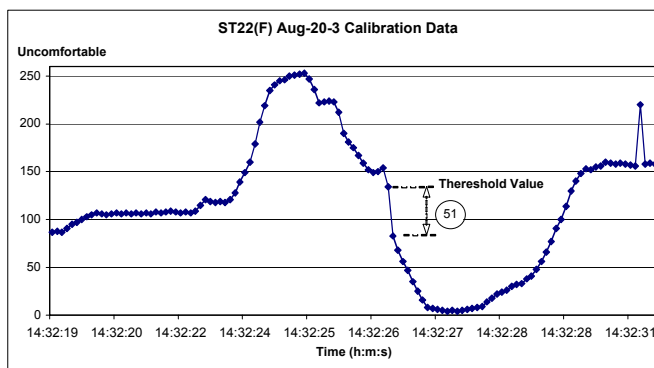


Fig. 4 Calibration data indicating the threshold value.

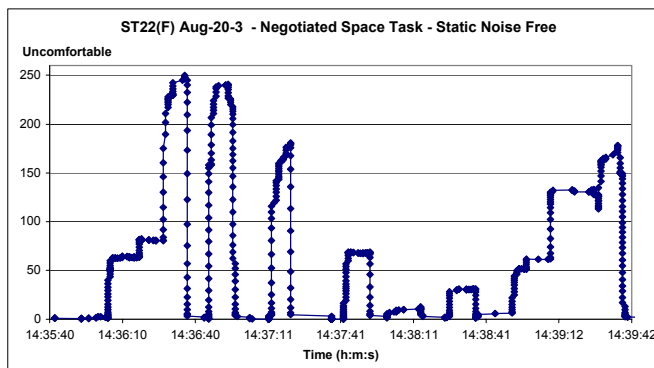


Fig. 5 Static free comfort level data after applying the filtering process using the threshold value shown in Fig. 4.

⁴ It is not our intention to make a contribution to the field of signal processing which has developed far more sophisticated techniques for noise filtering. Instead, we developed a simple technique that turned out to be sufficient for our particular application.

B. Analysis of Comfort Level Data

The comfort level data (e.g. Fig. 5) ranged from 0-255, proportional to the motion of the slider, with level 0 representing subjects' most comfortable state (i.e. corresponding to the position of the happy smiley face indicated on the device), while level 255 represents subjects' most uncomfortable states (i.e. sad smiley face).

The static free comfort level data of all 28 subjects were visually inspected and classified by the researchers. The data of seven subjects was considered to be very reliable: they clearly used the device consistently and the comfort data ranges from very comfortable to very uncomfortable, so we selected their comfort level data and video data for the present proof-of-concept analysis.

During the initial inspection of the comfort level data (backed by video observation), we found that the majority of the subjects forgot to use their Comfort Level Device after their first interaction task (i.e. after the Negotiated Space Task), see discussion section. For consistency, we decided in this study to concentrate only on the Negotiated Space Task. The fact that only some of the data was suitable for the analysis was not unexpected: a) this was the first time that the newly built device had been used in complex and live HRI trials, and b) this study was our first attempt to gain experience in difficult technical (e.g. interference) as well as methodological issues involved (e.g. how to remind subjects to use the device). We also expected from the outset that the device would only be suitable for particular tasks, we did not expect that the device could be applied generically across the range of all possible HRI scenarios.

For analysing the comfort data, we compared subjects' comfort level data with their corresponding behaviour shown in the experiment (recorded on video). We found that many of the recorded subjects' uncomfortable states corresponded to video sequences where subjects can be either seen moving the slider on the Comfort Level Device, or they were in a difficult situation such as crossing path with the robot, or the robot moving behind them while they were busy writing on the whiteboard. This suggests that a) subjects were willing and able to use the Comfort Level Device, at least in the Negotiated Space Task, and b) the comfort level data had not been produced randomly, but was correlated with subject's behaviour. These correspondences of video data and filtered comfort level data also indicated that the filtering process was successful in filtering out the noise while preserving the subjects' comfort profile recorded during the experiment. This confirms our first Research Question RQ1 - subjects did use the Comfort Level Device to indicate their discomfort. For future trials, it is intended to incorporate error checking and data verification into the RF data transfer link to the recording PC in order to further reduce problems with static.

IV. VIDEO ANALYSIS

By using the time stamps on the static free comfort data as a reference, we then matched the subjects' uncomfortable states with their video footages recorded during the experiments in order to determine exactly which types of robot behaviours caused the subjects to feel uncomfortable.

Figure 6, a, b, and c illustrate the first half of a video sequence where a subject and the robot crossed paths (the experimental design specifically encouraged such situations which are very common in human inhabited environments - so a robot should be able to deal with it). Here, the subject indicated her discomfort, through the Comfort Level Device, when the robot was heading towards her. The second half of the video sequence (Fig. 6, d, e and f) illustrates a situation where the subject immediately felt comfortable once she had finished crossing the robot's path. The second peak shown in Figure 6g illustrates the recorded subject's comfort level data for the situation shown in fig. 6 (a)-(f).

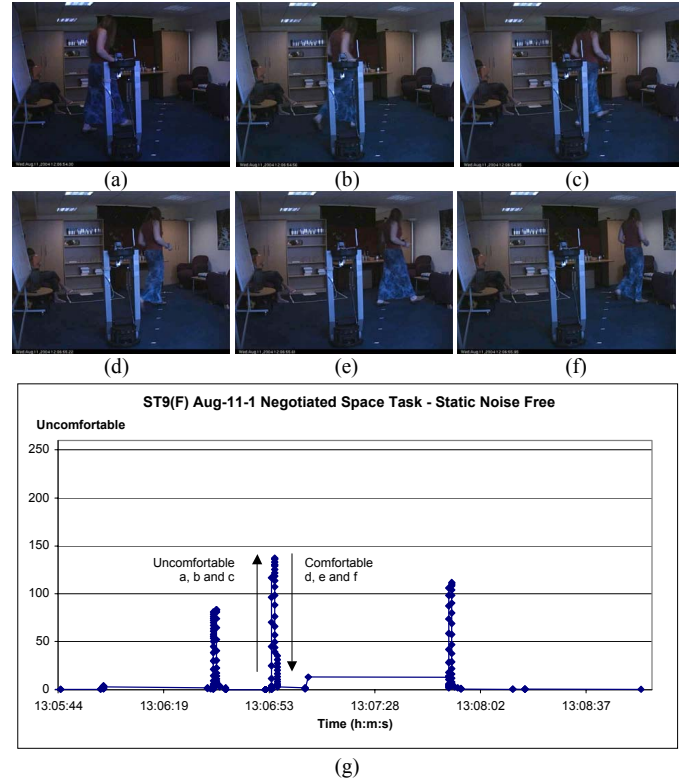


Fig. 6 Video sequences of a human-robot cross path scenario, where the robot stopped and said "after you" as soon as it detected the subject. (a)-(c) illustrate a scenario where the subject indicated that she was uncomfortable with the situation (see g). (d)-(f) illustrate the same scenario where the subject indicated she was comfortable (see g), (g) illustrates subject's comfort level during the Negotiated Space Task.

The comfort level data, along with video footage of all seven subjects revealed that in general there were 3 robot behaviours that were disliked by the majority of the subjects. Firstly, subjects do not like their path being blocked by the robot (Fig. 7a). Secondly, they also found it annoying when the robot moved behind them (Fig. 7b). This situation may be worsened by the robot's sonar sensors which were producing clicking noise that some subjects disliked (as indicated in the final questionnaires). Finally, subjects did not like the robot on a collision path heading toward them in a human-robot cross path scenario (Fig. 7c).

Two out of the seven subjects used only the Comfort Level Device to indicate their discomfort when the robot was

moving behind them (see Fig. 8). These subjects did not exhibit any other physical body language movements to indicate discomfort. This is in contrast to other subjects who used both the Comfort Level Device and body movements such as turning their head to glance at the robot, moving closer to the whiteboard to avoid collision, etc.

Based on our small sample size we cannot exclude the possibility that the discomfort signals in these situations were produced purely accidentally. However, the striking correspondence with situations where other subjects revealed discomfort, strongly suggests that the Comfort Level Device was used deliberately by the subjects to indicate discomfort. Thus, the Comfort Level Device was able to identify behaviours that are otherwise difficult to be noticed visually (i.e. *visually hidden uncomfortable states*) thus confirming RQ2.

One of the disadvantages of the Comfort Level Device is its sensitivity. We noticed that when the subject (see Fig. 9) opened the whiteboard pen cover, part of his arm motion was transferred to the comfort device slider through his index finger, hence the comfort level data registered ‘phantom data’ (i.e. registering the subject being in an uncomfortable/comfortable state).

V. CONCLUSIONS

In this paper we showed that the Comfort Level Device we developed, despite its limitations, was a useful tool that can be applied to the analysis of human-robot interaction, complementing other methods such as video analysis. The simple device turned out to be useful although a) the concept of ‘comfort’ was not specifically defined, and b) subjects had to ‘deliberately’ judge their comfort level and reflect their subjective comfort via explicit actions (manual movement of a slider). Before we began the trials it was unclear whether this extra cognitive, as well as manual ‘effort’ would be accepted by the subjects and yield useful results. However, our results show that the Comfort Level Device provided an insight and feedback from subjects’ point of views, revealing which of the robot’s behaviours subjects were uncomfortable with.

As expected for a first study using the device, a number of technical as well as methodological problems were identified. The device was suitable for one of the tasks/scenarios studied, but not the other (the majority of the subjects left their Comfort Level Device on the table throughout the Assistance Task). Generally, the device is likely to be more useful for some HRI tasks and contexts than for others. Note, we only reminded the subjects a couple of times during the Negotiated Space Task to use the device. It is thus not surprising that subjects then ignored the device in the second task. Whether the nature of the second task (sitting at a desk and writing) makes it unsuitable for the device, or the lack of reminders to use the device, needs to be investigated further.

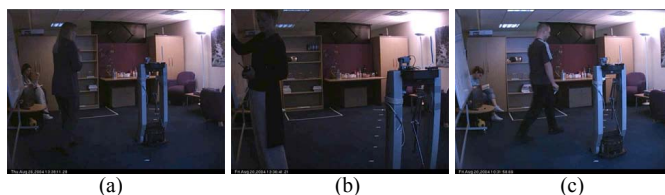


Fig. 7 Undesired robot behaviours, a) path blocked, b) robot behind subject, c) collision path.



Fig. 8 Visually hidden uncomfortable state, where subjects were feeling uncomfortable but continued writing on the whiteboard. This state was recorded and verified through video observation where subjects were seen moving the slider on the Comfort Level Device.



Fig. 9 Illustration of the phantom effect: uncomfortable data was recorded by the Comfort Level Device when a subject opened a whiteboard pen with both hands while still holding the Comfort Level Device in one hand.

Future work can investigate in more detail the suitability of the device for different scenarios, tasks, user groups⁵ etc. However, in this paper we provided proof-of-concept that the device was useful for the data analysis of seven subjects in the Negotiated Space task. Based on our results, it seems that the main issue regarding the Comfort Level Device is not to prove if it is useful or not (we have already shown its usefulness in certain cases), but to map out those HRI scenarios where it can make a significant contribution, in addition to improving on its usability and reliability. Where applicable, the device can replace or complement other devices for measuring subjects’ internal states.

Compared to our previous work, which relied solely on observational analysis [7,8], we consider the Comfort Level Device a useful tool. We provided proof-of-concept results for

⁵ For example, the device is likely not to be suitable for subjects with limitations in manual control or attention.

three selected robot behaviours that the majority of the subjects were uncomfortable with:

- a) Robot moving behind subject.
- b) Robot blocking subject's path.
- c) Robot on collision path with subject.

Subjects' preferred the robot not move behind them, not block their path and avoid being on a collision path (cross path scenario) with them. This situation often occurred when the robot made a turn in the area visibly labelled as 'robot only', leaving the rest of the simulated living room to the subjects. Subjects seemed to prefer the robot not to move around too much when it could interfere with subjects' movements. Also, they did not like to be interrupted in their activities or when the robot got in the subjects' way (i.e. created an obstruction) while subjects were busy with their tasks.

Care should be taken when analysing the Comfort Level Device data to avoid problems such as the phantom data caused by the movements that are caused as a side effect of the subjects' normal body movements and object manipulations.

In terms of our original research questions, we found:

- 1) A simple handheld device, such as our Comfort Level Device, does provide feedback on subjects' comfort level. We provided proof-of-concept data for seven subjects in the Negotiated Space Task.
- 2) We identified visually hidden uncomfortable states exhibited by 2 of the subjects which otherwise were very difficult to be identified, even by experienced video observers, without the help of the Comfort Level Device.

Further studies need to confirm the results in this paper using a larger sample size. Currently, we are correlating video data with comfort level data in greater detail in order to support and extend our findings in this paper. Furthermore, we will investigate ways to improve the Comfort Level Device to minimise static noise, reduce phantom data, and find ways to help subjects to continue remembering to use the Comfort Level Device. A very promising direction for future research concerns the possibility that the comfort level data, rather than just being used for post-experimental data analysis and interpretation only, could be used by the robot during the human-robot interaction trials to modify its behaviour style in order to adapt to subjects' preferences, likes and dislikes, an important prerequisite for a personalized robot companion [13].

ACKNOWLEDGEMENTS

The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020. The authors would like to thank Christina Kaouri, David Lee, Chrystopher Nehaniv, René te Boekhorst and Iain Werry for their contributions to the work. Four reviewers provided very valuable comments that helped us improving an earlier version of this paper.

REFERENCES

- [1] T. Saito, T. Shibata, K. Wada and K. Tanie, "Relationship between Interaction with the Mental Commit Robot and 'Change of Stress Reaction of the Elderly'", *Proc. IEEE CIRA*, pp. 16-20, 2003.
- [2] T. Kanda, H. Ishiguro, "Communication Robots for Elementary Schools", *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 54-63, April 2005.
- [3] B. Friedman, P. H. Kahn (Jr.), and J. Hagman, "Hardware Companions? – What online AIBO Discussion Forums Reveal about the Human-Robotic Relationship", *Proc. CHI'03 Conference on Digital Sociability*, vol. 5, pp. 273-279, 2003.
- [4] M. Walters, K. Dautenhahn, K. L. Koay, R. te Boekhorst, C. Nehaniv, I. Werry and D. Lee, "Close Encounters: Spatial distances between People and a Robot of Mechanistic Appearance", submitted for publication, July 2005.
- [5] K. Dautenhahn, S. Woods, C. Kaouri, M. Walters, K. L. Koay and I. Werry, "What is a Robot Companion – Friend, Assistant or Butler?", to appear in *Proc. IEEE IROS*, Aug 2-6 2005.
- [6] T. Salter, R. te Boekhorst, K. Dautenhahn, "Detecting and Analysing Children's Play Styles with Autonomous Mobile Robots: A Case Study Comparing Observational Data with Sensor Readings", *Proc. The 8th Conference on Intelligent Autonomous Systems (IAS-8)*, 10-13 March, Amsterdam, The Netherlands, IOS Press, pp. 61-70.
- [7] B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard, "Effects of repeated exposure to a humanoid robot on children with autism", *Proc. Universal Access and Assistive Technology (CWUAAT)*, Cambridge, UK, 22-24 March 2004, Springer Verlag (London), pp. 225-236.
- [8] K. Dautenhahn and I. Werry, "A Quantitative Technique for Analysing Robot-Human Interactions", *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1132-1138, 2002.
- [9] C. D. Kidd, and C. Breazeal, "Human-Robot Interaction Experiments: Lessons learned", *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 141-142, April 2005.
- [10] D. Mulsby, S. Greenberg and R Mander, "Prototyping an intelligent agent through Wizard of Oz", *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems*, Amsterdam, The Netherlands, ACM Press, pp. 277-284, 1993.
- [11] M. L. Walters, S. Woods, K. L. Koay and K. Dautenhahn, "Practical and Methodological Challenges in Designing and Conducting Human-Robot Interaction Studies", *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 110-120, April 2005.
- [12] COGNIRON. Website available at <http://www.cogniron.org>, 2005.
- [13] K. Dautenhahn, "Robots We Like to Live With?! - A Developmental Perspective on a Personalized, Life-Long Robot Companion", *Proc. IEEE Ro-man 2004*, Kurashiki, Japan, IEEE Press, pp. 17-22, 2004.

Empirical Results from Using a Comfort Level Device in Human-Robot Interaction Studies

K.L. Koay, K. Dautenhahn, S.N. Woods and M.L. Walters

University of Hertfordshire
School of Computer Science
College Lane
Hatfield, UK, AL10 9AB
+44 (0) 1707 28 {1026, 4333}

{K.L.Koay, K.Dautenhahn}@herts.ac.uk

ABSTRACT

This paper describes an extensive analysis of the comfort level data of 7 subjects with respect to 12 robot behaviours as part of a human-robot interaction trial. This includes robot action, proximity and motion relative to the subjects. Two researchers coded the video material, identifying visible states of discomfort displayed by subjects in relation to the robot's behaviour. Agreement between the coders varied from moderate to high, except for more ambiguous situations involving robot approach directions. The detected visible states of discomfort were correlated with the situations where the comfort level device (CLD) indicated states of discomfort. Results show that the uncomfortable states identified by both coders, and by either of the coders corresponded with 31% and 64% of the uncomfortable states identified by the subjects' CLD data (N=58), respectively. Conversely there was 72% agreement between subjects' CLD data and the uncomfortable states identified by both coders (N=25). Results show that the majority of the subjects expressed discomfort when the robot blocked their path or was on a collision course towards them, especially when the robot was within 3 meters proximity. Other observations include that the majority of subjects experienced discomfort when the robot was closer than 3m, within the social zone reserved for human-human face to face conversation, while they were performing a task. The advantages and disadvantages of the CLD in comparison to other techniques for assessing subjects' internal states are discussed and future work concludes the paper.

Categories and Subject Descriptors

A.m [Miscellaneous]: Human-Robot Interaction – *Social Robots*
I.2.m [Miscellaneous]: Robotics – *Mobile Robots*

General Terms

Measurement, Experimentation, Human Factors, Verification.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HRI'06, March 2–3, 2006, Salt Lake City, Utah, USA.
Copyright 2006 ACM 1-59593-294-1/06/0003...\$5.00.

Keywords

Human-Robot Interaction, Social Robot, Social Interaction, Comfort Level Device.

1. INTRODUCTION

Research in the field of Human-Robot Interaction (HRI) has grown over recent years, ranging from human development studies [8][9][15] to therapeutic robots [18][19][20] and robot companions [3][4][6][23][24]. Increasingly robots are able to operate in human-inhabited social environments. It is expected that the behaviour a robot exhibits will elicit certain human responses. Therefore it is not sufficient that future robot's behaviours are intrinsically safe (i.e. Asimov's Three Laws of Robotics [1]) but must also be socially acceptable. This has led to the issues of human social acceptance becoming a growing HRI research area using a human-centered perspective [6][9][19][23][24].

Our approach is to directly measure a subject's comfort level via a simple handheld device, where subjects use a continuous scale to judge their current comfort level throughout an HRI interaction trial. Two main strategies are commonly used for evaluating human-robot interaction from a human subjects' perspective 1) questionnaires e.g. as used in [4] and 2) video analysis of interactions, e.g. [2][11][18][20][21]. The latter is more appropriate for scenarios where verbal inquiry may not be possible (e.g. in the case of non-verbal subjects) [2][18], too intrusive, or might strongly bias the results [10]. In human-computer interaction and robotics, biofeedback sensors measuring physiological variables such as heart beat or skin conductance etc. have been investigated, e.g. [12][17]. However, the signal processing required for detecting affect and other internal states is often extensive, and sensors need to be attached to the subject. Deriving a high-level concept such as 'comfort' from rich physiological data is not straightforward, although subjects are very familiar with assessing their own subjective 'comfort level'.

In previous work [11] we discussed methodological issues of using a handheld device to assess subjects' subjective comfort levels in human-robot interaction experiments in a simulated living room scenario. We provided proof-of-concept showing that the processed data of the comfort level device can reveal certain situations where the subjects' felt uncomfortable in relation to three selected robot behaviours. We gave examples of situations indicating that the comfort level device can reveal uncomfortable states that were visually hidden, i.e. that cannot be identified by a

human observer on video recordings of the interactions. In the present work we did an extensive analysis of the comfort level data of the 7 subjects (whose comfort level data were considered to be very reliable cf. [11]) for the Negotiated Space Task in order to validate the CLD. First, we analysed the comfort level data with respect to 12 robot behaviours, including robot action, proximity and motion relative to subjects. Second, two researchers coded the video material, identifying visible states of discomfort in relation to the robot's behaviour. Agreement between the coders varied from moderate to high except for more ambiguous situations involving robot approach directions. The detected visible states of discomfort were correlated with the situations where the comfort level device (CLD) indicated states of discomfort.

The findings of robot behaviours where subjects experienced discomfort are discussed, followed by the limitations of the CLD, its advantages and disadvantages in comparison to other techniques for assessing subjects' internal states. An outline of future work concludes the paper.

2. HUMAN-ROBOT INTERACTION STUDY

The exploratory study involved single human subjects in a simulated living room scenario. It was carried out at the University of Hertfordshire between July and August 2004. This study was conducted using a commercially available, human-scaled, PeopleBot™ robot. The main aim of the study was to evaluate, in a task oriented living room scenario, different social behaviour and interaction styles of the PeopleBot™ robot from a human-centred perspective. A sample of adult volunteers was recruited from the University of Hertfordshire, balanced for gender, background, and familiarity with technology. All subjects completed consent forms and were not paid for participation.

2.1 Experimental Setup – The Simulated Living Room

The original room measured 8.5 x 4.75m and was partitioned off at one end to form an area that served as a control area for the Wizard-of-Oz [14][22] operators and provided space for the control, network and recording equipment. Two researchers controlled the robot: one person was driving the robot, the other one was controlling the robot's speech output and the robot's camera movements [22]. Decisions on the robot's movements were based on views from the robot's onboard camera and surrounding cameras located in the room. The room was decorated as a living room in order to provide a comfortable environment relevant to the COGNIRON¹ project which studies a robot companion in domestic scenarios.

2.2 The Handheld Comfort Level Device

We built a handheld comfort level monitoring device that would allow subjects to indicate their internal comfort level during the experiment (see Figure 1).

The device uses a slider control, located at one edge of the box, to receive subjects' comfort level feedback. The slider can be moved

easily by the subjects using either a thumb or finger to indicate their comfort level. The slider scale was marked on one end of the slider with a happy face, to indicate the subject was comfortable with the robot's behaviour, and a sad face on the other end, to indicate discomfort with the robot's behaviour. The device used a 2.4GHz radio signal data link to send numbers representing the slider position to a PC mounted receiver, which recorded the slider position approximately 10 times per second. The data was time stamped and saved in a file for later synchronisation and analysis in conjunction with the video material. Figure 2 illustrates a subject using the handheld comfort level device to indicate her comfort/discomfort with the robot behaviours.



Figure 1. Photograph of the handheld Comfort Level Device.



Figure 2. A subject using the handheld Comfort Level Device to indicate her discomfort with the robot's behaviour.

2.3 The Experimental Procedure

The experiment was supervised by an experimenter who introduced and explained the trials to the subject. Each subject spent approximately 50 minutes in the simulated living room with only the robot and the experimenter present who interfered as little as possible with the trials. The following phases of the experimental procedure are relevant to the present paper.

Introduction: A general welcome phase where the robot was introduced to the subject when they entered the simulated living room. An information sheet was given to the subject to read along with a consent form to be signed, then questionnaires were completed. The robot moved around the room whilst the subject completed these initial questionnaires in order to familiarize the subject with the robot.

Comfort Level Device: Before subjects proceeded to the main trial, they were given a Comfort Level Device (Figure 1) and were asked to try it out and operate it a few times (for calibration purposes and in order to provide an opportunity for the subject to get accustomed to the device²). Next, they were told to use it

¹ <http://www.cogniron.org>

² The handheld device might possibly provide an additional potential source of discomfort. By allowing time for the subject to get used to the device we tried to reduce this effect. Any such additional discomfort is likely to persist during the whole trial and is less likely to influence the changes in the levels of comfort/discomfort which were our primary concern. Focusing on changes in the comfort levels has a second advantage: it makes the data more independent of any 'moods' that a

throughout the main trial to indicate their comfort level during the trial. The terms comfort/discomfort were not defined prior to the experiment, and was left to the subjects' self-interpretation. Therefore the term 'uncomfortable' in the context of this study can be considered as any negative feeling that subjects felt due to a situation they were in that could occur during any intended/unintended interaction with the robot.

A subset of the data collected in this way during the trials forms the basis of this paper.

Main Trial: The main trial consisted of two tasks, a Negotiated Space Task and an Assistance Task. The Negotiated Space Task involved the robot moving in the room while the subject went through a pile of books placed on the table, remembering one title at a time, walking over and writing down each title on the whiteboard. The Assistance Task involved the subject sitting at the table, copying the book titles from the whiteboard onto a piece of paper and underlining specific letters with a red/highlighter pen. The robot was responsible for bringing the missing red/highlighter pen to the table. The two tasks were chosen as they match two key scenarios studied in the COGNIRON project. At the end of these two task scenarios, the subject completed questionnaire. The Main Trial was then repeated³.

Final Phase: The final phase involved the subjects completing several questionnaires. The analysis of the questionnaires and other data collected during the trials is reported in different publications e.g. [4,23].

3. VIDEO ANALYSIS

Video analysis has been used widely in the field of studying human interaction. Its history can be traced back to the early 1940s, where David Efron [5] in his pioneering study of gestures traced his subjects' hands and arms movement through projection of motion-picture films of his subjects frame by frame onto graph paper. The motivation for using video analysis in this study was to identify subjects' **Instances of Discomfort** (IoDs) and

particular subject might be in e.g. on a particular day, assuming that such moods are persistent over a longer period of time. Further studies need to clarify this issue.

³ In terms of the experimental design of our study it might be interesting to see how subjects in a control group, not using the handheld device, would behave. However, the main purpose of our study was to identify whether the handheld device could be used to relate subjects' subjective judgments of comfort/discomfort with observable behaviour. A group of subjects using other, more sophisticated and expensive (e.g. physiological) devices to identify discomfort could serve as a suitable control group. However, it is not clear how to easily deduce comfort/discomfort from physiological data. Asking for vocalisations (e.g. "I don't feel comfortable now", or verbal ratings on a scale from one to ten) did not seem appropriate either since it would have interfered with the reading/writing tasks that the subjects were performing during the experiment. Also, moving a slider with one finger seemed easier to use compared to the effort required to pinpoint verbally exact moments of discomfort. Vocalizations would also not be able to provide fine graded quantitative data. Note, our primary aim is to develop a reliable Comfort Level Device for human-robot trials. Thus, a control group involving human-human interaction, instead of human-robot interaction, did not seem suitable either. Our main motivation was to use a simple, very inexpensive device, that can easily be replicated by any talented person with certain engineering skills, using a simple data analysis technique, and validate the approach in HRI experiments.

associated robot behaviours that caused discomfort during the HRI trials. An annotation tool [13] was used to perform a quantitative evaluation of observational data (video footage of the experimental trials) by first identifying subjects' IoDs, followed by the associated robot behaviours of each IoD. This coding process was performed manually on a second-by-second basis. Subjects' IoDs were identified based on body language, body movements, facial expressions or utterances that indicated discomfort according to the judgment of the coders. This includes e.g. jumpy or jerky body movements, surprised facial expressions and intermittent checking where the robot is.

Details of robot behaviours of interest are discussed in the next subsection.

Table 1. Video Annotation Coding Scheme

Behaviour Code	Action	Robot
A1	Robot Moving behind Subject	Moving behind the subject
A2	Robot Blocking Subject's Path or On Collision Course	<ul style="list-style-type: none"> • In subject's way or restricting the subject from moving freely • Moving straight at constant speed regardless of subject in its path • Pause forward motion • Change path or break stride across the room
A3	Robot is rotating on the Spot	Rotates in 'robot only area' (a specifically marked area in the experimental room) or area in front of desk
A4	Robot is avoiding subject or avoids getting too close to subject	In the process of avoiding subject (either responding/taking initiative)
A5	Robot is observing subject	Camera pointing ('looking'), tracking or searching for subject
A6	Others	Any observed behaviour that does not fit into any of the behaviour categories defined above
Proximity		
P1	Close	Less than 1 m
P2	Intermediate	1 to 3 m
P3	Far	More than 3 m
Motion		
M1	Approach	Moving towards subject's position
M2	Receding	Moving away from subject's position
M3	Stop	Not moving

Note, some of the robot behaviours may occur one after another, but only the first robot behaviour that influenced a particular state of subject's discomfort was coded.

3.1 Video Coding Scheme

The coding scheme for the 12 robot behaviours was produced to help coders identify robot behaviours that subjects were uncomfortable with. This included the robot's actions, as well as proximity and motion relative to the subjects. The 12 behaviours in the coding scheme cover 6 different robot actions, 3 different robot proximities and 3 different robot motions. The 6 different actions and 3 different motions were derived from the robot

behaviours identified from the recorded video footage of the experimental trials. The 3 different proximities (distance between subject and robot) used were inspired by Hall's [7] work on proxemics⁴. A detailed description of all the 12 robot behaviours is shown in Table 1.

3.2 Video Coding Methods

Two coders were asked to independently code the video annotation using the pre-defined coding scheme to determine the reliability of the coding scheme and identifiable behaviours. They were told to code the robot's behaviour when they noticed the subject exhibiting any signs of discomfort from the video footage. Discomfort in this context refers to the subjects' feeling negative due to any intended or unintended interaction with the robot.

The video annotation process involved the coder going through video footage for 7 subjects who consistently used the CLD, cf. discussion of methodological issues in [11]. The resolution used for annotating the video footage was set at 1 possible IoD per second. For each identified IoD, the coders were required to score the robot's behaviour according to robot action, robot proximity and robot motion relative to the subject, based on the video coding scheme discussed above. Before the actual video coding process began, both video coders were asked to familiarise themselves with the coding scheme using a test video sample.

3.3 Inter-rater Reliability Test (Cohen's kappa)

To determine the consistency of the robot behaviours coded by both independent raters, the two raters' results were compared using a statistical test (Cohen's kappa). From the independent coding of the videos for 7 subjects, coder C_a recorded 52 IoDs and C_b recorded 35 IoDs. Of all the identified IoDs, 25 present in both coders' annotated video data.

These 25 IoDs were used as a basis for the inter-rater reliability tests of agreement for the robot behaviours. The inter-rater agreement between the two independent coders was calculated using the Kappa Coefficient statistics and revealed the following: **Action** – A1, kappa=.364, p=.041 with overall agreement of 72%; A2, kappa=.59, p=.003 with overall agreement of 80%; A6, kappa=.468, p=.006 with overall agreement of 92%; **Proximity** – P1, kappa=.595, p=0.003 with overall agreement of 88%; P2, kappa=.565, p=.004 with overall agreement of 84%; P3, kappa=.648, p=.001 with overall agreement of 96%; **Motion** – M1, kappa=.606, p=.002 with overall agreement of 84%; M2, kappa=.606, p=.002 with overall agreement of 84%.

The majority of the results represent moderate to high agreement with the exception of **Action** – A1 which had a low to moderate inter-rater reliability score. Further analysis revealed that this was mostly caused by ambiguous situations where it was very difficult for the video coder to decide whether the robot's action falls under A1 or A2 (e.g. visually one can see the robot was moving behind the subject, but was the subject's discomfort due to the fact that robot was moving behind her or because the robot was restricting her from moving freely?).

The kappa statistics for **Action** – A3, A5, and **Motion** – M3 were not computed because no such behaviours were recorded by both coders within the 25 matching IoD samples. Kappa statistics was not computed for **Action** – A4 as a symmetric 2-way table in which the values of the first variable match the values of the second variable was not present for this behaviour (within the 25 matching IoD samples).

4. COMFORT LEVEL DATA ANALYSIS

This section discusses the results of the CLD in terms of three different analyses. Firstly, subjects' subjective IoDs (subsection 4.1), and secondly the correlation of subjects' subjective IoDs with their respective IoDs identified by video coders (subsection 4.2). This includes the correlation of subjects' subjective IoDs with their respective IoDs identified by both video coders under the strict-matching rule (subsection 4.2.1), and with either one of the video coders under the relaxed-matching rule (subsection 4.2.2.) The correlation results discussed in subsection 4.2 lead to the discussion of the issues of visually hidden IoDs (subsection 4.3) and phantom IoDs (subsection 4.4).

4.1 Subjects' Comfort Level Data

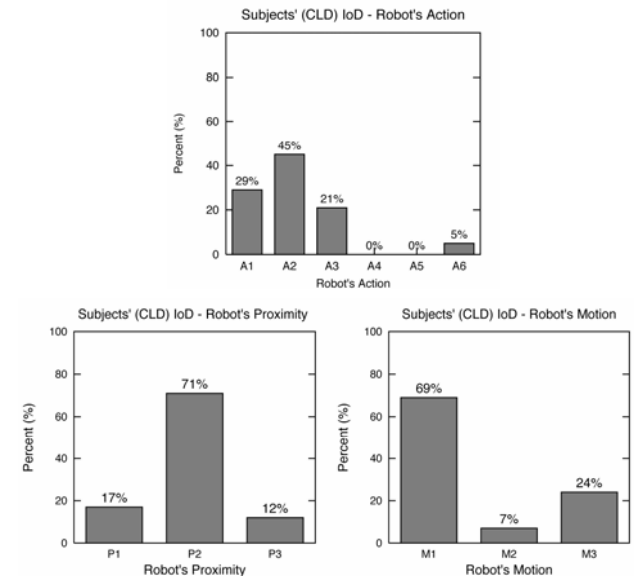


Figure 3. Subject's subjective video annotation results of robot behaviours that subjects were uncomfortable with.

A total of 66 IoDs were obtained from the 7 subjects usage of the CLD. Eight of these were identified as phantom IoDs through the process of video annotation. As discussed in [11] phantom data occurs whenever the CLD's slider is unintentionally moved caused by subjects' body movements, see subsection 4.4. The IoDs in the comfort level data were identified using the techniques described in detail in [11].

The results of the observational data for the robot behaviours based on subjects' comfort level data, excluding the phantom IoDs, revealed that the majority of subjects indicated they were uncomfortable with the robot's action A2 (45%, N=58), the robot's proximity P2 (71%, N=58) around the subjects and robot motion M1 (69%, N=58).

⁴ Due to practical reasons (reliability of judging distances from video footage) we could not follow *exactly* Hall's social spaces zones.

4.2 Video Annotation of Subjects' IoDs and Associated Robot Behaviours

In the first phase of the video annotation process, coders C_a and C_b identified IoDs in the video footage of 7 subjects. Findings revealed that the number of IoDs observed by both coders were different: C_a recorded a total of 52 IoDs while C_b recorded a total of 35 IoDs. The breakdown of each coder's observed IoDs for each subject is shown in Table 2.

Table 2. Summary of Subjects' IoD

		Number of Subjects' IoDs						
		S_a	S_b	S_c	S_d	S_e	S_f	S_g
Video Coder	C_a	6	2	5	15	20	2	2
	C_b	9	1	1	10	10	2	2
Subjects	CLD	14	3	3	12	29	3	2
Matching rules	$C_a \cap CLD$	3	2	3	6	11	0	1
	$C_b \cap CLD$	6	1	1	6	8	0	1
	$C_a \cap C_b$	5	0	1	7	9	2	1
	$(C_a \cap C_b) \cap CLD$	3	0	1	5	8	0	1
	$(C_a \cup C_b) \cap CLD$	6	3	3	9	15	0	1
Video Length (min.)		6:32	3:34	3:39	4:29	4:49	4:11	3:14

The difference in the number of visible IoDs identified by both coders was not unexpected, as the detection of subjects' uncomfortable behaviours based on visual cues is a very difficult task, even for trained coders. Factors such as physical restrictions or disabilities may effect the subjects' body movement [16], and can lead to misinterpretation. Also, due to the nature of our unconstrained human-robot interaction experiments, subjects of different heights were allowed to move freely within a reasonably large experimental area to complete a task. Therefore, it was very difficult to have an ideal camera view on the subjects throughout the experiment, despite the use of several video cameras. Note, that the coding of subjects' IoDs may be subject to bias (i.e. coder dependent), even though a pilot test was first introduced before the main video annotation to overcome/minimize the problem of coder dependency. It was also expected that there would be a low percentage of agreement between the matched IoDs for both coders, and the CLD generated IoDs.

Therefore, two matching rules were introduced and applied to the IoDs identified by the coders, and the CLD generated IoDs to assess how they correlated. The two selected rules were 1) Strict-matching Rule – the IoDs identified by both coders agreed with the CLD generated IoDs i.e. $(C_a \cap C_b) \cap CLD$, and 2) Relaxed-matching Rule – the IoDs identified by either of the coders agreed with the CLD generated IoDs i.e. $(C_a \cup C_b) \cap CLD$.

The correlation result from the strict-matching rule was as expected low, scoring just 18 IoDs as opposed to the relaxed-matching rule which scored a total of 37 IoDs, nearly double of the strict-matching rule's score (see Table 2). The robot behaviours that caused these IoDs are discussed in subsection 4.2.1 for the strict-matching rule and subsection 4.2.2 for the relaxed-matching rule.

4.2.1 Strict-Matching Rule of Subjects' Comfort Level based on Video Annotation

Overall, 31% (N=58) of the IoDs where subjects indicated they were uncomfortable (during the experiment, through the CLD) matched the video annotation results from both video coders according to the strict matching rule $(C_a \cap C_b) \cap CLD$.

Based on the strict-matching rule between the video observation results annotated by the video coders and the subjects' CLD data, we find 72% (N=18x3) of matching robot behaviours out of all behaviours coded from the video footage. Of these matching behaviours, A2 scored the highest in the robot action category with 67% (N=15), P2 scored the highest in robot proximity category with 92% (N=12) and M1 scored the highest in robot motion category with 100% (N=12). Table 2 shows IoDs for all seven subjects S_i as generated by the CLD and scored by the two different coders C_a and C_b .

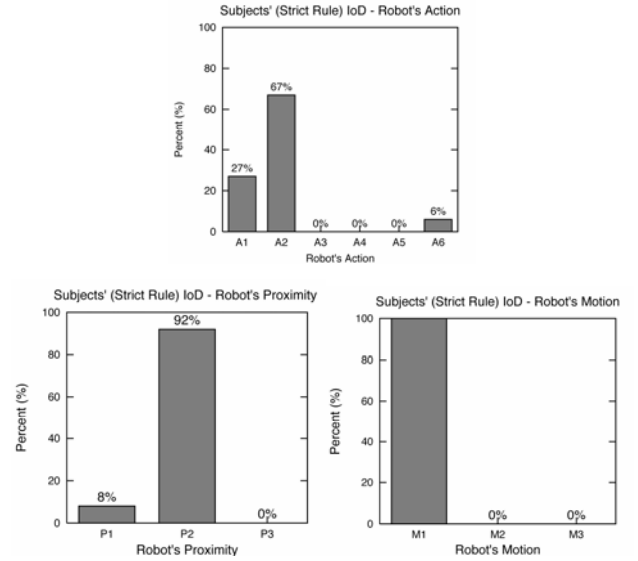
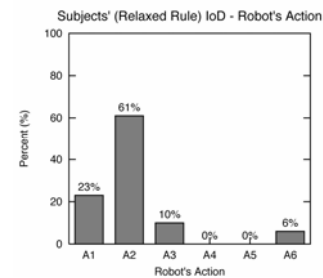


Figure 4. Strict-matching rule video annotation results of robot behaviours that subjects were uncomfortable with.

4.2.2 Relaxed-Matching Rule of Subjects' Comfort Level based on Video Annotation

Overall 64% (N=58) of the IoDs where subjects indicated they were uncomfortable (during the experiment, through the CLD) matched the video annotation results from either of the video coders according to the relaxed-matching rule $(C_a \cup C_b) \cap CLD$.

Based on the relaxed-matching rule between the video observation results annotated by the video coders and the subjects' CLD data, we find 79% (N=37x3) of matching robot behaviours out of all behaviours coded from the video footage. Of these matching behaviours, A2 scored the highest in the robot action category with 61% (N=31), P2 scored the highest in robot distance category with 69% (N=29) and M1 scored the highest in robot motion category with 82% (N=28).



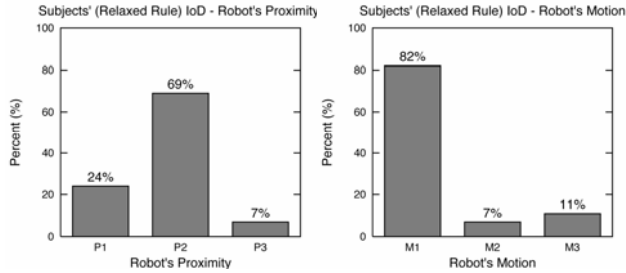


Figure 5. Relaxed-matching rule video annotation results of robot behaviours that subjects were uncomfortable with.

4.3 Visually Hidden IoDs

As shown in Figure 6, a total of 29 visually hidden IoDs were identified from the subjects' CLD data (i.e. IoD's indicated by the subjects via the CLD, but not identified by the video coders). Results from the video analysis on visually hidden IoDs show that 8 of the 29 visually hidden IoDs were actually phantom IoDs (more will be discussed in subsection 4.4). The remaining 21 visually hidden IoDs were observed in four subjects S_a , S_d , S_e and S_f with 8, 3, 7, and 3 IoDs respectively. These visually hidden IoDs suggest three possible explanations: I) Subjects did not always exhibit any visible physical body language movement to indicate their discomfort, II) Subjects did exhibit physical body movements to indicate their discomfort but these were too subtle to be identified by the video coders, and III) due to the nature of the video data in a large unconstrained experimental area, subjects might have been partially or completely out of the cameras' visual field.

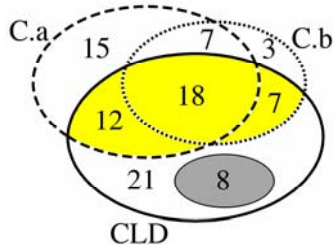


Figure 6. A general overview of the IoDs coded by both coders (i.e. C_a and C_b) and those indicated by the subjects themselves (i.e. CLD). The figure also shows that there were 21 visually hidden IoDs and 8 phantom IoDs.

4.4 Phantom IoDs

Of the 8 phantom IoDs identified, all of them were caused by the subjects' index finger/thumb's unintentional motion resulting from opening and closing the whiteboard pen cover. Seven of the phantom IoDs were identified in the video footage of subject S_e while the remaining 1 IoD was identified in the video footage of subject S_g .

The results from the video observation revealed that the configuration of subjects' grasping/holding the CLD played a pivotal role in the number of generated phantom IoDs (see Figure 7). This was clearly demonstrated for subject S_e , where the majority of the phantom IoDs (i.e. 7 IoDs) came from. Subject S_e was holding the CLD in a configuration where he used his index finger for adjusting the CLD's slider control (i.e. slider control facing the subject's index finger), whereas the other 6 subjects held the CLD in a configuration where they could easily access

the CLD's slider control with their thumb (i.e. slider control facing the subjects' thumb) which only generated 1 phantom IoD.

Therefore, it is clear that the problem of phantom IoD we identified in this study could be minimised by advising subjects to grasp the CLD in a particular configuration where the slider control is facing her thumb.



Figure 7. Phantom IoD mainly resulted from the way the subject was holding the CLD.

5. DISCUSSION

The results of the 3 different video analyses discussed in subsection 4.1 and subsection 4.2 revealed 3 different ways of viewing the results of the robot behaviours subjects were uncomfortable with. "Subjects' IoD" which were shown in Figure 3 represents more detailed results of robot behaviours subjects were uncomfortable with based on subjects' CLD data. "Strict-matching rule" which was shown in Figure 4 represents the main robot behaviours subjects were most uncomfortable with when comparing IoD's identified in the video footage and CLD data. "Relaxed-matching rule" which was shown in Figure 5 represents robot behaviours subjects were uncomfortable with including a bigger variety of robot behaviours due to the relaxed-matching rule.

Results show that the uncomfortable states identified by both or either coders corresponded with 31% and 64% of the uncomfortable states identified by the subjects' CLD data ($N=58$), respectively. Conversely there was 72% agreement between subjects' CLD data and the uncomfortable states identified by both coders ($N=25$). This was a good result considering coders C_a and C_b coded a total of 52 and 35 IoDs respectively (which was only 6 IoDs less than the subjects' CLD IoDs) for all 7 subjects. This was only a very small subset of all possible 1828 IoDs. This number is equivalent to the total number of time coding steps (that each coder annotated) for all the experiments.

The overall video analysis results of the CLD data from all 3 different analyses (Figure 3, Figure 4, and Figure 5) shows a similar profile and ranking orders of robot behaviours subjects were uncomfortable with. It is difficult to conclude in this paper which robot action subjects were most uncomfortable with, as the live and unconstrained HRI trial does not permit us to design predefined experimental conditions for robot behaviour e.g. with an equal number of robot behaviours in each category. This limits the identification of robot behaviours subjects were most uncomfortable with. Nevertheless, the results give an insight into the dominant robot behaviour from each category, as well as showing which robot behaviours subjects were uncomfortable with in the context of our live and unconstrained HRI trials. As shown in all the 3 different analyses of robot behaviours that subjects were uncomfortable with (see section 4), the dominant robot behaviour for the category Action, Proximity and Motion

were A2 (Robot blocking Subject's Path or On Collision Course), P2 (Intermediate proximity – 1 to 3 meters) and M1 (Approach) respectively. The coded robot behaviours (i.e. robot action, proximity and motion) for each associated IoDs were interrelated, but were coded separately into 3 categories. Therefore, the main finding can be interpreted as that a majority of subjects were discomforted when the robot blocked their path or was on collision course towards them; especially when the robot was within 3 m.

The current results confirm previous reported findings [11] (i.e. subjects dislike the robot moving behind them, blocking their path or on collision path toward them.). However, the current study identified that subjects were also uncomfortable with robot action A3 (Robot is rotating on the spot: in an area marked as 'robot only area') and A6 (Others: actions that were not defined in the coding scheme which included: 1) when the robot said "excuse me" or "after you" which could interrupt the subject in his task, or the subject did not understand what the robot was saying; 2) the robot was heading towards a possible collision with an obstacle.)

A minority of subjects were uncomfortable even when the robot was 3 meters or more away from them. This could be due to the subject being nervous with the robot moving within their workspace. The majority of subjects experienced discomfort when the robot was closer than 3m. This was understandable, as 3 m is within Hall's [7] Social Zone, which is mainly reserved for face to face conversation. It was shown in our previous work [24] that human-robot approach distances can be comparable to human-human social distances in some cases. Therefore it is possible that subjects may feel insecure, threatened or intruded upon when a robot enters their personal or social zones.

Results for robot motion indicated that the majority of subjects were uncomfortable when the robot was approaching them. Especially when they were writing on the whiteboard (i.e. robot was moving behind them), or trying to move across the experimental area between the whiteboard and the desk, where the books were located (see Figure 2).

Of the 6 robot actions that were defined in the video coding scheme, only one subject was not comfortable with robot action A3. This may be due to the subject feeling insecure having the robot perform an action behind or out of subject's visual field (i.e. video observation revealed that the subject was seen occasionally looking at the robot when the robot was performing action A3.) We did not find any indication that the subjects were annoyed by robot's action A4 (robot is avoiding subject or avoids getting too close to subject) nor A5 (robot is observing subject). This may not come as a surprise as action A4 may be difficult to be identified by the video coders as this action often followed after action A1 or A2. Therefore subjects' IoDs may have already been associated to either action A1 or A2. Whereas, action A5 was not associated with any of subjects' IoDs, this may be due to the fact that subjects were concentrating on their tasks and did not notice the robot's camera, which was quite small in relation to the robot's physical size, that was observing/tracking them.

In this paper we have shown the usefulness, reliability, and practicality of the CLD, and we believe it is a useful tool to facilitate subjects' comfort levels in live and unconstrained HRI trials where subjects' mobility is an essential part of the experiment. We have also shown that the CLD is useful for post-

experimental data analysis as the CLD data was able to reveal visually hidden IoDs that were otherwise difficult to be identified by the video coders. The device not only allows the detection of both visible and hidden IoDs, but also answer the problem of video analysis where: (1) subjects might have been partially or completely out of the cameras' field of view, and (2) subjects are out of focus or far behind in the background of the camera's field of view (i.e. poor resolution on the focused subjects).

The CLD allows direct identification of subjects IoDs. This is the most crucial and difficult part of video analysis. Even with experienced video coders, there is no guarantee that they will observe all relevant behaviours related to subjects' IoDs. These IoDs are typically revealed through body language or subtle cues such as facial expressions or utterances.

A limitation of using the CLD is that subjects might forget to use the device. Also, the device may not be useful in experiments where subjects are required to use both their hands to complete their task(s).

Comparing to other biofeedback devices such as galvanic skin response sensors attached to the finger, while they may interfere less with the manual task, stress levels caused by the task and other environmental factors may interfere with the results.

We have highlighted in this paper how the identified phantom IoD problem can be overcome. Like most newly developed technology there is still room for improvement and we expect other unforeseen problems/bugs (e.g. different phantom IoD problem) to surface in our future experiments, which will allow us to improve the CLD.

Future work needs to address the issues of encouraging subjects to use the CLD continuously, not just for extreme behaviours. In future it would be useful if the image of a dimensioned rectangular grid could be superimposed on the experimental floor area in the video recording. This would allow video coders to estimate the robot's proximity to subjects accurately. As this grid would be purely a video artifact there would be no chance of influencing subjects' movement during the experiment. A laser scanner might be used instead. More trials need to be carried out with the CLD to identify and overcome any hidden limitations. While in this study we used the CLD for post data analysis, a very promising direction for the future use of the CLD is to allow subjects to modify the attributes of a robot's behaviours style in order to adapt to subjects' preferences, likes and dislikes as a requirement for developing a personalized robot companion [3].

6. ACKNOWLEDGEMENTS

The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

7. REFERENCES

- [1] Asimov, I. I, Robot (a collection of short stories originally published between 1940 and 1950), Grafton Books, London, 1968.
- [2] Dautenhahn, K., and Werry, I. A Quantitative Technique for Analysing Robot-Human Interactions", Proceedings of IEEE

- IROS (Lausanne Switzerland, September -October 2002), 1132-1138.
- [3] Dautenhahn, K. Robots We Like to Live With?! – A Developmental Perspective on a Personalized, Life-Long Robot Companion. Proceedings of IEEE RO-MAN (Kurashiki Okayama, September 2004), 17-22.
- [4] Dautenhahn, K., Woods, S., Kaouri, C., Walters, M., Koay, K.L., and Werry, I. What is a Robot Companion - Friend, Assistant or Butler? Proceedings of IEEE IROS (2005), 1488-1493.
- [5] Efron, D. *Gesture, Race, and Culture*. The Hague: Mouton, 1972. (Originally published as *Gesture and Environment*. New York: King's Crown Press, 1941.)
- [6] Friedman, B., Kahn (Jr.), P.H., and Hagman, J. Hardware Companions? – What online AIBO Discussion Forums Reveal about the Human-Robotic Relationship, Proceedings of CHI'03 (Fort Lauderdale Florida, April 2003), vol. 5, 273-279.
- [7] Hall, E.T. *The Hidden Dimension: Man's Use of Space in Public and Private*. The Bodley Head Ltd, London UK, 1966.
- [8] Kanda, T., Hirano, T, Eaton, D., and Ishiguro, H. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human Computer Interaction (Special issue on human-robot interaction)*, 19, 1-2 (2004), 61-84.
- [9] Kanda, T., and Ishiguro, H. Communication Robots for Elementary Schools. Proceedings of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (Hatfield Hertfordshire, April 2005), 54-63.
- [10] Kidd, C.D., and Breazeal, C. Human-Robot Interaction Experiments: Lessons learned. Proceeding of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (Hatfield Hertfordshire, April 2005), 141-142.
- [11] Koay, K.L., Walters, M.L., and Dautenhahn, K. Methodological Issues using a Comfort Level Device in Human-Robot Interactions. Proceedings of IEEE RO-MAN 2005 (Nashville TN, August 2005), 359-364.
- [12] Kulic, D., and Croft, E. Estimating intent for human-robot interaction. Proceedings of IEEE Int. Conf. on Advanced Robotics (2003), 810-815.
- [13] Loehr, D., and Harper, L. Commonplace Tools for Studying Commonplace Interactions: Practitioners' Notes on Entry-Level Video Analysis. *Visual Communication*, 2, 2 (2003), 225-233.
- [14] Maulsby, D., Greenberg, S., and Mander, R. Prototyping an intelligent agent through Wizard of Oz. Proceedings of CHI'93 (Amsterdam, 1993), ACM Press, 277-284.
- [15] Michaud, F., Laplante, J-F., Larouche, H., Duquette, A., Caron, S., Letourneau, D., and Masson, P. Autonomous Spherical Mobile Robot for Child-Development Studies. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 35, 4 (July 2005), 471-480.
- [16] Pease, A. *Body Language: How to Read Others' Thoughts by Their Gestures (Overcoming Common Problems)*. Sheldon Press, 1997.
- [17] Rani, R., Sarkar, N., Smith, C., and Kirby, L. Anxiety detecting robotic systems – Towards implicit human-robot collaboration. *Robotica*, 22, 1 (2004), 85-95.
- [18] Robins, B., Dautenhahn, K., te Boekhorst, R., and Billard, A. Effects of repeated exposure to a humanoid robot on children with autism. Proceeding of Universal Access and Assistive Technology (Cambridge UK, March 2004), 225-236.
- [19] Saito, T., Shibata, T., Wada, K., and Tanie, K. Relationship between Interaction with the Mental Commit Robot and 'Change of Stress Reaction of the Elderly. Proceeding of IEEE CIRA (Kobe Japan, July 2003), 16-20.
- [20] Salter, T., te Boekhorst, R., and Dautenhahn, K. Detecting and Analysing Children's Play Styles with Autonomous Mobile Robots: A Case Study Comparing Observational Data with Sensor Readings. Proceedings of the IAS-8 (Amsterdam The Netherlands, March 2004), 61-70.
- [21] Tardif, C., Plumet, M.-H., Beaudichon, J., Waller, D., Bouvard, M., and Leboyer, M. Micro-analysis of social interactions between autistic children and normal adults in semi-structured play situations. *Int. J. Behav. Develop.*, 18, 4 (1995), 727-747.
- [22] Walters, M.L., Woods, S., Koay, K.L., and Dautenhahn, K. Practical and Methodological Challenges in Designing and Conducting Human-Robot Interaction Studies. Proceedings of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (Hatfield, April 2005), 110-120.
- [23] Walters, M.L., Dautenhahn, K., te Boekhorst, R., Koay, K.L., Kaouri, C, Woods, S., Nehaniv, C., Lee, D., and Werry, I. The Influence of Subjects' Personality Traits on Personal Spatial Zones in a Human-Robot Interaction Experiment. Proceedings of IEEE RO-MAN (Nashville TN, August 2005), 347-352.
- [24] Walters, M., Dautenhahn, K., Koay, K.L., Kaouri, K., te Boekhorst, R., Nehaniv, C., Werry, I., and Lee, D. Close Encounters: Spatial distances between People and a Robot of Mechanistic Appearance. Accepted for *Humanoids 200*

A Flexible Task Knowledge Representation for Service Robots

Steffen Knoop, Sven R. Schmidt-Rohr, and Rüdiger Dillmann

IAIM, Institute of Computer Science and Engineering (CSE),
University of Karlsruhe(TH), Germany
{knoop,srsr,dillmann}@ira.uka.de

Abstract. Robots that are designed to act in human-centered environments put up the need for a flexible and adaptive representation of task knowledge. This results on the one hand from the continuously changing and hardly predictable state of an environment that is populated with humans and robots. On the other hand, a task knowledge description of a robot which cooperates with humans has to be adaptable and extendable.

This paper presents a task knowledge representation for service robots called *Flexible Programs (FP)* and the environment for execution of FPs. Flexible Programs can be created manually, or by using the results of machine learning approaches like *Programming by Demonstration*. It is possible to change, adapt and extend this task knowledge at runtime.

1 Introduction

Service robots which possess manipulation capabilities require a powerful task representation. Such a representation has to provide several features, e.g. representations for goals, tasks, possibly conditions, etc. Representing tasks has been studied for quite a long time in the domain of task planning, and each execution environment which is able to perform a sequence of actions is equipped with a (implicit or explicit) task representation. A simple task could be the goal of navigating to a target position. Here, the task and the goal both are implicitly given by the target position and the algorithms, which are part of the robot model.

More complex tasks which include manipulation, variables within the task, and a larger set of actions and exceptions, have to be modeled more explicitly. Especially the capability of generating executable tasks from observation of human demonstrators (*Programming by Demonstration*, see [1]) and the demand of legibility for humans require for an explicit task model. This paper describes the task model and execution environment which is used at IAIM for the service robot *AlbertII*, see fig. 1. The robot is equipped with a differential drive platform, a 7dof manipulator, a three-fingered hand, stereo camera, speech in/out and a touch display on the front. The target application for this robot is the assistance for humans in household environments.

2 State of the Art

Within the field of robotics, almost as many task descriptions as robots can be found. Each robot which performs complex tasks requires a task description, but no standard task description has evolved yet. This mainly results from the fact that the task knowledge is very often closely linked to the dedicated execution system, and thus not very portable. Nevertheless, some task descriptions have been developed during the last decade which are highly interesting and have proved their reliability and effectiveness in several projects, mainly evolving from the field of robot programming languages.

True robot programming languages are mostly built as extensions to standard programming languages like C (see Colbert [2]) or LISP (see RAPs [3][4], Frob [5]). Thus, runtime extension of tasks is not possible or only according to predefined rules. A survey can be found in [5]. Declarative task descriptions often rely on static task descriptions. These tree-like (e.g. TDL [6], or Hierarchical Task Networks [7], [8]) or procedural (see [9], [10]) descriptions can be extracted from user demonstrations and are closely related to PbD output action sequences. However, most existing declarative task descriptions have been designed for manual task compilation and do not provide extensive support for condition management, which is a prominent requirement for the execution of learned action sequences.

3 Framework and Software Architecture

A software architecture that is able to control a mobile robot basically has to cover three aspects to comprise all capabilities of the system: Control of the hardware, representation of the environment and integration of knowledge about skill and task execution. Including the demand for a lifelong learning system, these requirements have to be fulfilled in a way that the system remains extensible.

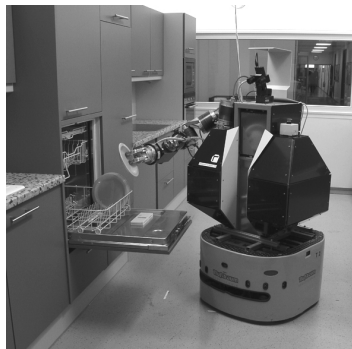


Fig. 1. Robot AlbertII in kitchen environment

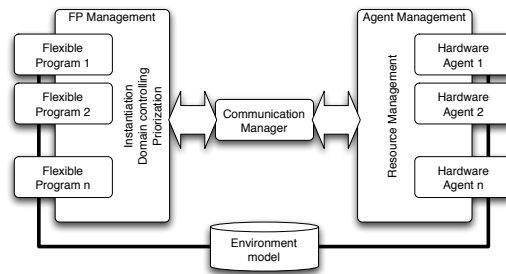


Fig. 2. CORBA-based software architecture with flexible programs, hardware agents and communication scheme

An expedient approach to design such an architecture is therefore to identify the functional components as hardware abstraction, environment model and skill and task execution. Our software architecture consists consequently of the following components (see figure 2):

The *hardware agents* encapsulate all hardware specific functions. There is a hardware agent for each hardware in the robot system, e.g. one for the robot arm, one for the hand and one for the voice. Skill and task knowledge is represented by *Flexible Programs*. Example skills are "drive to position" and "open door", tasks can be "transport an object" or "put object on table". All information about the environment is stored in an *environment model*. This can hold all kinds of data: Coordinates, objects, relations, features, images or sounds. A communication bus connects all these components.

The components are in detail:

- **The communication infrastructure** consists of a notification distribution instance, where clients can subscribe for certain notification types. Notifications may be delivered by internal or external sources.
- **Hardware agents** (resources) represent real or virtual sensors and/or actuators. There is an agent for *laser scanner* and for *cameras*, but there may be also an agent for *detection of humans* which incorporates laser scanner and vision information.
- **The agent manager** administrates all hardware agents and provides the resource management. Each notification that is passed to an agent is filtered by the agent manager. Thus, unauthorized commands (from instances which have not locked the called resource) can be intercepted.
- **Flexible programs (FPs)** contain the skill and task knowledge. The FPs are the core of the robot control architecture. Learning, within our context, means creation, extension and adaption of flexible programs.
- **Domain controlling and supervision** as well as FP instantiation and priority control is done by the flexible program supervisor. Depending on the current context, FPs are created, prioritized or deleted. In later development, learning capabilities will be incorporated in this component.
- **The environment model** holds environmental data as well as the robot's internal state. It is implemented as a blackboard. An intelligent xml data base, which is being developed at our research group, has been integrated for storage of task, skill and object feature knowledge.

The implementation of the proposed software architecture (see figure 2) consists of a set of CORBA object types, which communicate using a publish-subscribe mechanism (see also [11]). Each public object's interface is described in CORBA's Interface Definition Language, which enables the use of different programming languages within the same framework. A full description of the execution environment, including the robot hardware and the execution framework, can be found in [12].

4 Task knowledge representation

Within our concept, the flexible programs hold the task knowledge. This knowledge is represented as the number, type, parameterization and order of the basic skills and other flexible programs used. In addition to all generic requirements for programming languages, it is important that the task description is adaptive and extendable at runtime. This is a major requirements for integration of online learning and adaptation methods, which is an essential requirement for service robots in human-centered environments. Thus, the language must not need a compilation process, a generated or adapted program should be directly executable.

Many robot programming languages have been developed throughout the last years, some of which are very mature and used in many applications (e.g. *ESL*, see [13], or *TDL*, see [14]). But in particular the program extendibility at runtime has not been an explicit demand.

Therefore, we have developed a modular and extensible task description called *Flexible Programs (FP)*. The tasks described as FPs are based on a set of primitive actions and functionality to combine and select different paths through these actions.

4.1 Flexible Programs

Similar to TDL and HTN planning (see section 2), the task is described as a hierarchical network which is processed with a depth-first search strategy. In contrary to other representations, the FP task tree is not statically built, but instantiated dynamically at runtime. This section describes the tree nodes and tree instantiation.

One node is described by a parameter set $\mathcal{P} = \{Id, C_{pre}, C_{post}, C_{rt}, R, S, P, A\}$ with *Id* a unique identifier, C_{pre} , C_{post} and C_{rt} a set of pre-, post- and runtime-conditions, *R* a rating function, *S* a success measure, *P* the child prospect and *A* the action. Two different node types exist:

- *Leaves* contain actions or other flexible programs which have to be executed. Their prospect is always empty $P = \emptyset$.
- *Inner nodes* contain execution rules which control the order and execution of child nodes within the prospect *P* which must not be empty. Their action *A* is always $A = \emptyset$.

Consequently, all robot actions are contained in tree leaves which are connected by inner branching nodes.

The child prospect in branching nodes contains possible child configurations. It is composed of a set of $n > 0$ *seats* with $m > 0$ *candidates* each. The seats are arranged in $p \leq n$ *parallel groups*. Tree expansion then is done by dynamically selecting one candidate for each seat at runtime. The state of a node can be *virtual* (no children selected yet), *visited* (children selected and instantiated) or *completed* (execution of all children finished).

Candidate selection is done based on two criteria:

- The preconditions P_{pre} are evaluated to determine whether a candidate node can be executed. Only if the given preconditions are *true*, the candidate can be chosen.
- In the case where more than one candidate may be chosen for the same seat based on the precondition evaluation, each candidate’s rating function R is evaluated. This function can be any kind of function based on accessible environment parameters (e.g. time of day, light conditions, etc.). The candidate with highest rating will be chosen.

Candidate selection is thus highly depending on environmental and robot parameters. To select the right subnode for a given situation, the seats are filled only when the parent node is under execution for the first time. This means that selection of actions can be done based on the same situation in which it will be executed.

Children which are arranged in seats within *parallel groups* in the prospect will be executed in parallel. The children are entered simultaneously, and the parallel group is treated in the superior context equal to one node, which means that all dependent actions are entered when all children of the parallel group are *completed*. This implies that parallel seats must be arranged consecutively in the prospect. It is important to note that parallel subtrees can not be evaluated for colliding actions. So it is in the responsibility of the FP builder to avoid collisions in parallel subtasks.

As stated before, tree leaves contain basic actions. This can either be a command to one of the hardware abstractions, or it can be the execution command and parameters for another flexible program. In both cases, the action will be marked as *completed* when a corresponding notification from the target instance returns. An example for dynamic candidate selection and tree expansion can

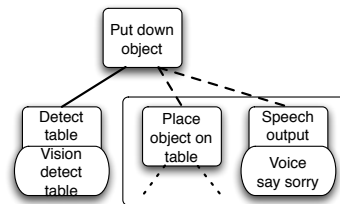


Fig. 3. Runtime candidate selection based on world and robot state with two candidates for the following action

be seen in fig. 3. Based on the result of the first action, the second one can be chosen. Therefore, the postcondition of the first action has to be part of the precondition of the second.

Tree leaves trigger primitive actions within the hardware abstraction layer. These actions can be further parameterized from the environment model by

using variables from the global database. It can be specified in each node which slot has to be filled from which variable. In the same way, action results (return values) can be stored in the global database, thus giving the possibility to pass parameters from one node to another.

Flexible program format and generation: Flexible programs are specified and stored completely in XML. An interpreter engine loads and executes the FPs according to external or internal trigger events. This approach provides a clear

```

<entity name="REPutObjToTable" type="node">
  <constraintbox><constraints type="start"> <equal>
    <arg>holdObject</arg>
    <arg const="true">true</arg>
  </equal>
  </constraints>
  <constraints type="goal" (...) />
</constraintbox>
<action datatype="keyvalue" />
<prospect seats="2"><seat>
  <candidate>
    <name>REMoveToTable</name>
    <unit>0</unit> <inunit>0</inunit>
  </candidate>
</seat>
<seat>
  <candidate>
    <name>REPlaceObj</name>
    <unit>1</unit> <inunit>0</inunit>
  </candidate>
</seat>
</prospect>
</entity>

```

Fig. 4. Flexible Program node specification example in XML notation. Some fields are left out for clarity of presentation

separation of task description and execution, which increases portability between different robots and keeps the task knowledge base open-ended. Subtasks can easily be reused in different Flexible Programs, in the same way as primitive actions are included in each FP. An example xml is shown in fig. 4.

The task knowledge of a robot in human environments will always be incomplete due to the changing environment. Thus, it is important that the robot can learn new tasks from a user. This problem is being addressed within the framework of *Programming by Demonstration* (see [15],[16]). The output of a PbD learning cycle is a task description which is independent from the target robot: It only includes the segmented actions performed by the human demonstrator. The resulting task description is called *Macro-Operator* (MO). It is now

possible to compile such a MO into a Flexible Program. To accomplish this, additional knowledge about the target system is required and has to be added and included in the task description. This work currently being addressed to gain (semi-) automatic robot program creation.

Of course, it is possible to create and edit Flexible Programs manually by compiling actions, alternatives, and conditions into an execution tree. This process is simplified by making reuse of existing task knowledge.

5 Experiments and Results

The FP description and interpreter module have been used and tested in several applications with the service robot AlbertII. Two examples are shown: A "pick" (fig. 5) and a complex "pick and place" task (fig. 7).

The first one demonstrates the candidate selection and parameter passing capabilities with a real robot system: Depending on the result of the object localization *Detect object*, the hand preshape, TCP position, approach direction and grasp type are chosen. Depending on the result of the grasp action, one of two speech outputs is selected. Finally, platform and arm retreat in parallel to a safe position. The execution can be seen in fig. 6.

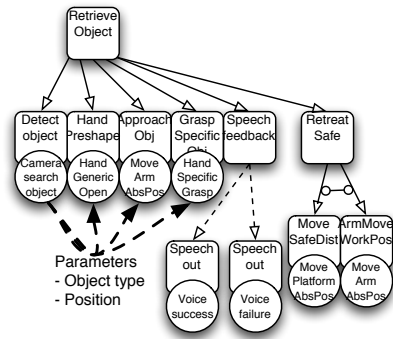


Fig. 5. Flexible program *Retrieve Object*. Processed in depth-first search

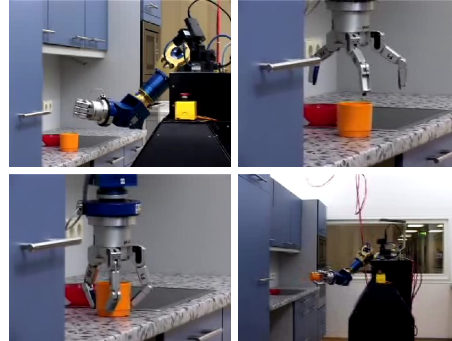


Fig. 6. Top left: Detect object, top right: Preshape hand, bottom left: Grasp object, bottom right: Retreat

A complex example FP is shown in figure 7. The depicted task moves the robot in front of the fridge, takes an object out of it, including the *open* and *close* of the fridge, and places the object on a given table. It has been executed successfully in simulation.

In the given fully instantiated (i.e. no candidates, execution finished) example, several mechanisms can be found. Leaves, i.e. actions, are marked with a circle, inner nodes, containing tree functionality, are depicted as squares. Parallel actions are linked together.

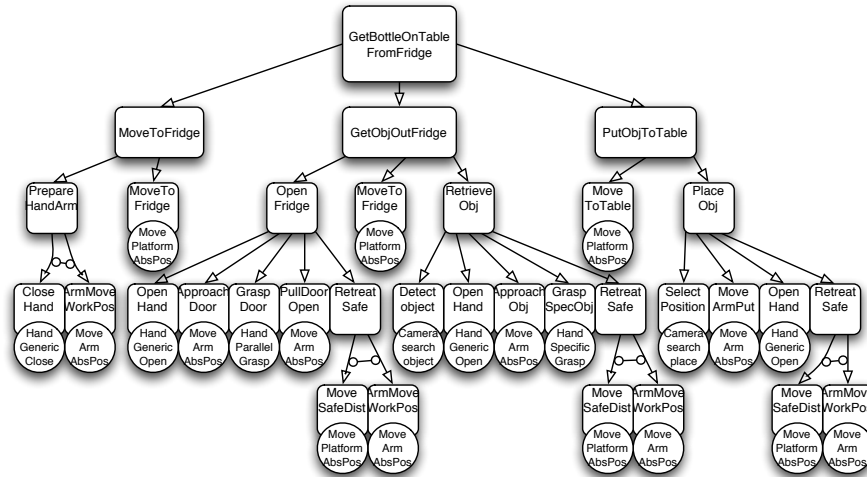


Fig. 7. Complex task example as Flexible Program

- Sub-FPs are reused several times, see e.g. the *retreat safe* block.
- Some non-colliding children are executed in parallel, see e.g. *Close hand* and *ArmMove WorkPos* in *PrepareHandArm*.

6 Conclusion

In this paper, we described a task representation for service robots which is set up of parameterized basic actions. These actions are linked within the task description in a tree like manner, and the tree is read and executed in a depth-first search. The task representation is able to cope with exchange of variables, and program parts can be reused in other situations and contexts. The tree-like structure also supports this reusability.

The presented approach of modeling task knowledge has proven to work on a real robot, giving the possibility to automatically choose the best alternative solution for each subtask at runtime. The design of the task knowledge representation further allows for task adaptation and extension at runtime. This will be the next step in the development of Flexible Programs.

7 Acknowledgements

The work described in this paper was partially conducted within the EU Integrated Project COGNIRON (“The Cognitive Companion”) and funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

References

1. M. Ehrenmann, R. Zöllner, O. Rogalla, S. Vacek, and R. Dillmann, "Observation in programming by demonstration: Training and execution environment," in *IEEE International Conference on Humanoid Robots*, Karlsruhe, Germany, Octobre 2003.
2. K. Konolige, K. L. Myers, E. H. Ruspini, and A. Saffiotti, "The Saphira architecture: A design for autonomy," *Journal of experimental & theoretical artificial intelligence: JETAI*, vol. 9, no. 1, pp. 215–235, 1997.
3. R. J. Firby, "The rap language manual, version 2," I/NET Inc., Tech. Rep., 2001.
4. R. Firby, "Adaptive execution in complex dynamic worlds," Ph.D. dissertation, 1989.
5. G. Hager, J. Peterson, and P. Hudak, "A language for declarative robotic programming," in *Proceedings of the International Conference on Robotics and Automation*, 1999, pp. 1144–1151.
6. R. Simmons and D. Apfelbaum, "A task description language for robot control," in *Proceedings of the Conference on Intelligent Robots and Systems (IROS)*, 1998.
7. K. Erol, J. Hendler, and D. S. Nau, "HTN planning: Complexity and expressivity," in *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, vol. 2. Seattle, Washington, USA: AAAI Press/MIT Press, 1994, pp. 1123–1128.
8. K. Erol, J. Hendler, and D. Nau, "Semantics for HTN planning," Computer Science Dept., University of Maryland, Tech. Rep. CS-TR-3239, 1994. [Online]. Available: citeseer.ist.psu.edu/erol94semantics.html
9. K. L. Myers, "A procedural knowledge approach to task-level control," in *Proceedings of the 3rd International Conference on Artificial Intelligence Planning Systems (AIPS-96)*, B. Drabble, Ed. AAAI Press, 1996, pp. 158–165.
10. B. Jensen, G. Froidevaux, X. Greppin, A. Lorotte, L. Mayor, M. Meisser, G. Ramel, and R. Siegwart, "The interactive autonomous mobile system robot," in *Proceedings of the 2002 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*. IEEE, 2002, pp. 1221–1227, ePFL, Lausanne, Switzerland.
11. R. Dillmann, "Interactive natural programming of robots: Introductory overview," in *DREH 2002*. Tsukuba, Ibaraki, Japan: Tsukuba Research Center, AIST, December 2002.
12. S. Knoop, S. Vacek, R. Zöllner, C. Au, and R. Dillmann., "A corba-based distributed software architecture for control of service robots," in *Proceedings of the IEEE International Conference on Intelligent Robots and System*. Sendai, Japan: IEEE Institute of Electrical and Electronics Engineers, 2004, pp. 3656–3661.
13. R. P. Bonasso, J. Firby, E. Gat, D. Kortenkamp, D. P. Miller, and M. G. Slack, "Experiences with an architecture for intelligent, reactive agents," vol. 9, no. 2/3, 1997, pp. 237–256. [Online]. Available: citeseer.ist.psu.edu/bonasso97experiences.html
14. R. Simmons and D. Apfelbaum, "A task description language for robot control," 1998. [Online]. Available: citeseer.ist.psu.edu/simmons98task.html
15. M. Ehrenmann, O. Rogalla, R. Zöllner, and R. Dillmann, "Teaching service robots complex tasks: Programming by demonstration for workshop and household environments," in *Proceedings of the 2001 International Conference on Field and Service Robots (FSR)*, vol. 1, Helsinki, Finland, 11.–13. Juni 2001, pp. 397–402.
16. R. Dillmann, O. Rogalla, M. Ehrenmann, R. Zöllner, and M. Bordegoni, "Learning robot behaviour and skills based on human demonstration and advice: the machine learning paradigm," in *9th International Symposium of Robotics Research (ISRR '99)*, Snowbird, Utah, USA, 9.-12. Oktober 1999, pp. 229–238.

Human-style interaction with a robot for cooperative learning of scene objects^{*}

Shuyin Li, Axel Haasch, Britta Wrede, Jannik Fritsch, Gerhard Sagerer
Faculty of Technology
Bielefeld University
33594 Bielefeld, Germany
{shuyinli, ahaasch, bwrede, jannik, sagerer}@techfak.uni-bielefeld.de

ABSTRACT

In research on human-robot interaction the interest is currently shifting from uni-modal dialog systems to multi-modal interaction schemes. We present a system for human-style interaction with a robot that is integrated on our mobile robot BIRON. To model the dialog we adopt an extended grounding concept with a mechanism to handle multi-modal in- and output where object references are resolved by the interaction with an object attention system (OAS). The OAS integrates multiple input from, e.g., the object and gesture recognition systems and provides the information for a common representation. This representation can be accessed by both modules and combines symbolic verbal attributes with sensor-based features. We argue that such a representation is necessary to achieve a robust and efficient information processing.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Natural language*; I.4.8 [Image Processing and Computer Vision]: Scene Analysis—*Tracking, Object recognition*; H.2.5 [Heterogeneous Databases]: Heterogeneous Databases

General Terms

Management

1. INTRODUCTION

Multi-modality is one of the most important features that characterize human-human social interaction. Based on this

^{*}This work has been partially supported by the European Union within the 'Cognitive Robot Companion' (COGN-IRON) project (FP6-002020) and by the German Research Foundation within the Collaborative Research Center 'Situated Artificial Communicators' as well as the Graduate Program 'Task Oriented Communication'.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICMI'05, October 4–6, 2005, Trento, Italy.
Copyright 2005 ACM 1-59593-028-0/05/0010 ...\$5.00.

observation, designers of computer systems have been trying to integrate multi-modal input and output mechanisms in human-machine interactions to enable more intuitive operations for human users. Given the huge amount of existing multi-modal systems with quite different notions of multi-modality we first want to discuss two important aspects of modality intuitiveness to clarify our position. In general it is assumed that the degree of intuitiveness of a modality will determine the smoothness and efficiency of the interaction it facilitates. Thus, the question is what modalities are the most intuitive for users? We argue that the answer is of evolutionary nature and is application-dependent. For example, the mouse has become the major modality in human-computer interaction for many users. For these users the mouse is probably one of the most intuitive ways to operate a computer although it has little to do with the natural communication channels they use in human-human communication such as speech. In contrast, many other users prefer to write commands directly. We can imagine that future generations will find other modalities more intuitive than a mouse. The feeling of intuitiveness in computer operation is thus continuously changing. Even the same people may judge the intuitiveness of one modality differently when operating different computer systems. For example, people tend to anthropomorphize mobile robots when interacting with them. We argue that this is even more the case for application areas where a robot is supposed to assist and accompany humans in a social environment such as private households. Our envisioned robot is supposed to "live" in a private household. Our long term goal, therefore, is to endow it with social capabilities so that it can become some sort of "companion". This means that we do not envision it to serve humans in a master-slave manner as people tend to suppose. Rather, a robot companion should *cooperate* with humans to achieve certain goals. One basic function of such a Robot Companion is to be able to learn interactively about its environment. We therefore devised the *home-tour* scenario within our project where a human user is supposed to show his/her home to a newly purchased robot.

Based on our goal to design a robot that can be accepted by the user as a companion we argue that the interaction with such machines should be in a human style. We define the term *human style modalities* as *multi-modal communication channels that humans are biologically equipped for and (learn to) use from their birth*. Typical examples are speech and gestures. These modalities differ from other modalities like mouse and keyboard in that they are learned nat-

urally and without the use of artificial devices. In contrast, we define artificial modalities that are commonly used for human-computer interaction as *virtual modalities* because their effect is a virtual one which is only observable by humans via an artificial interface (e.g., the computer display). Thus, since people tend to anthropomorphize robots by expecting human-like abilities and attributes we conclude that robots should be endowed with human-style modalities for interaction.

A further aspect concerns the knowledge representation. Consider our home-tour scenario where the interaction will involve a high degree of deictic activity as the user will point to diverse objects in the environment. Thus, when the user points to his/her computer and explains “This is my computer”, the robot should be able to recognize the user’s gesture, find the computer and associate the symbolic name “computer” with a visual representation. This knowledge should be stored in a multi-modal way in order to be retrievable from different modules for further interactions. Psychological theories of knowledge representation in humans suggest that the symbolic name of an object is associated with its sensory features like its image or haptic characteristics (e.g. [3]). When activating the name of an object other features of the object are also activated. This indicates that the cognitive representation of objects in humans is multi-modal and therefore allows for multi-modal processing of information. In order for a robot to cooperate with a human, it should therefore be able to also process multi-modal information to build a representation similar to that of its human communication partner to support a better mutual understanding. We therefore developed a multi-modal representation scheme.

To summarize our position shortly: The intuitiveness of modalities needed for operation of computers and machines has an evolutionary aspect and depends on the individual applications. In our Robot Companion domain we are interested in human-style multi-modality that should be considered for both, communication channels and the representation of knowledge. Its impact is of functional and technical nature.

In this paper we will first present the multi-modal processing strategies of the Dialog System (section 4) and the Object Attention System (section 5) followed by a detailed description of our multi-modal representation scheme in section 6. Results in the form of a dialog example will be given in section 7.

2. RELATED WORK

While there is an increasing interest in multi-modal interfaces there is only a very limited number of applications that use human-style modalities based on an integrated multi-modal knowledge representation.

That multi-modal cues are beneficial for increasing the robustness in human-robot interaction has been shown for example in [13] where communication errors are detected by using not only speech recognition scores but also by including laser data to infer the presence or absence of communication partners and noise sources. Repair actions also involve the use of multiple modalities by either driving around to actively search for a communication partner or by offering buttons as alternative communication channels in noisy environments. More human-style interactions have been suggested in [1] by modeling a *naturalness support behavior*. This be-

havior includes verbal strategies by inserting filler phrases, as well as non-verbal reactions such as nodding or head turning as reactions to environmental noise. The authors report positive reactions by the users but extensive evaluations still remain to be done. The benefits of using multi-modalities in a learning scenario have been demonstrated in several applications. For example, the robot Leonardo [2] can learn the names of buttons when a human communication partner points them out by verbal and deictic instructions. Leonardo also learns specific interactions with these buttons by demonstration. However, while the interactive capabilities of Leonardo are quite realistic the underlying representations are simple and no new objects can be learned. An impressive system running on a mobile robot is presented by Ghidary et al. [6]. The robot is able to learn objects by analyzing speech commands and hand postures of the user. The user gives verbal information about the object’s size and can describe the spatial relations between objects, e.g., by phrases like ‘left of my hand’. The rectangular views of the learned objects are stored in a map representing the robot’s environment and can be used for later interactions. Although the interaction system is very limited and the resulting map is rather coarse, this system can be compared to our approach as it also builds up a long-term memory about objects in the environment. However, we focus on a more detailed representation of objects and their later recognition in order to support natural human-robot interaction going beyond simple navigational tasks.

In general there is a tendency to either focus on building a robust representation while neglecting interaction smoothness or vice versa. We argue that in order to build a robot that is able to be perceived as a companion it needs both, a more natural interaction based on an integrated multi-modal representation.

3. OVERALL SYSTEM

The scene acquisition system described in this paper is being implemented on our mobile robot BIRON. BIRON’s hardware platform is a Pioneer PeopleBot from ActivMedia with a pan-tilt camera for face tracking and object and gesture recognition, stereo microphones for speaker localization and speech recognition, and a SICK laser range finder for locating legs of potential communication partners. The overall architecture [10] of BIRON is based on a hybrid control mechanism and has three layers: a reactive, an intermediate and a deliberative layer (see Fig. 1). Modules that are responsible for reactive feedback of the system are set on the reactive layer: the *Person Attention System* detects potential communication partners and the OAS detects objects that users refer to. Since these are purely data-driven processes they belong to the reactive layer. Modules responsible for higher-level processing that involve top-down, expectation-driven strategies such as a planner or the Dialog System, are located on the deliberative layer. The Scene Model, which contains a multi-modal representation of the objects that the system has observed and can be seen as an intermediary between the Dialog System and the OAS, is consequently located on the intermediate layer. The communication between modules is carried out via XCF (XML Enabled Communication Framework). The system is centrally controlled by the so-called *Execution Supervisor* on the intermediate layer. It coordinates the module operations and makes sure that neither the reactive layer mod-

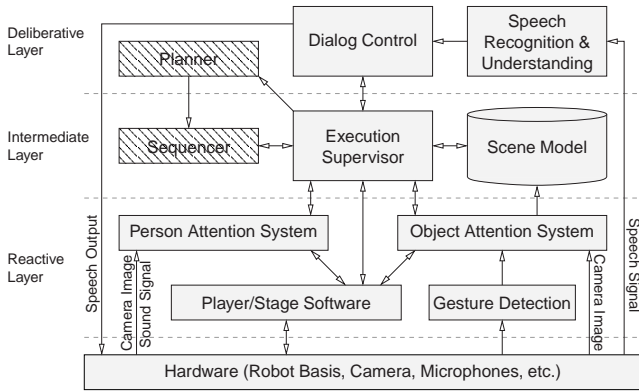


Figure 1: Overview of the BIRON architecture (optional modules currently not implemented are drawn in grey).

ules control the deliberative layer modules nor vice versa. Instead, it exerts control by taking into account the overall system state. This architecture allows for both fast reaction to dynamic environmental changes and extensive high-level planning and reasoning activities.

Since we focus on the interaction capabilities of BIRON, we do currently not integrate a planner and a sequencer. The Person Attention System establishes the basis for the interaction with users but is not involved in the process of resolving object references. In the following we therefore only describe the interfacing between the Dialog System, the OAS and the Scene Model.

4. THE DIALOG SYSTEM

The dialog system of BIRON is responsible for carrying out interactions with the user including handling miscommunications [16], guiding the discourse, and transferring user utterances to internal command for the robot control system to execute tasks. A dialog is made up of contributions from the dialog partners. Two central questions of dialog modeling are therefore (1) how to represent individual contributions represented and (2) how to represent the dynamic change of the dialog state represented which is triggered by individual contributions successively. In subsection 4.1 and 4.2 we are going to present our answers to these two questions. In section 4.3 we focus on the mechanism that the implemented dialog model provides for the integration of speech and visual input. A more detailed account on this integration will be given in section 5.

4.1 The structure of a contribution

Conversants contribute to a dialog in a multi-modal way. McNeill [12] investigated the relationship between speech and simultaneous conversational gesture and claims that the production of them are motivated by one single semantic source, the so-called “idea unit”. Inspired by this finding, we represent the conversants’ contribution as the so-called “interaction unit” that includes two important stages of the language production process. The structure of the interaction unit is illustrated in Fig. 2. An interaction unit has two layers: a *domain layer* and a *conversation layer*. The *domain layer* mirrors the cognitive activities of a dialog participant that motivate language production: If the interaction unit represents an utterance to be produced by the

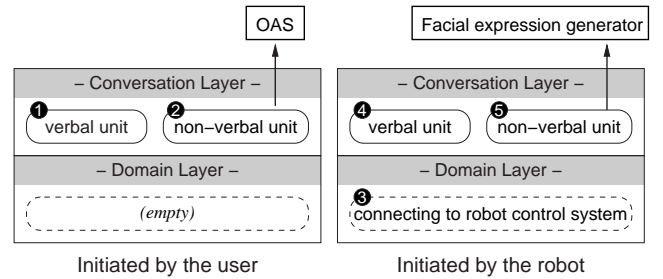


Figure 2: Processing flow in interaction units

robot itself the domain layer is where the Dialog System accesses the robot’s control system or knowledge base. If the interaction unit represents an utterance of the user the domain layer remains empty because we do not make any assumptions about the user’s cognitive activities behind the language front. In future work this may be replaced by a user model. The *conversation layer* transforms the intention that is created based on the results of these cognitive activities to language. For example, based on a successful follow behavior of the robot that is reported to the domain layer by the robot control system, the conversation layer formulates and synthesizes a message such as “OK, I follow you”. The conversation layer consists of two units: a verbal and a non-verbal unit. Based on the intention that results from the domain layer they are responsible for generating output in verbal and non-verbal way, respectively.

The precondition of language production is successful language perception. Before reacting, i.e., before creating his/her own interaction unit to produce a contribution, a conversant first needs to understand the semantic meaning of his/her dialog partner’s contribution by studying the verbal and non-verbal unit on the conversation layer of the dialog partner’s interaction unit. Therefore the processing in the interaction unit should start from this language perception phase. Figure 2 illustrates how the robot processes user contributions. In our system, the user’s verbal information as delivered by the Speech Understanding System initiates the creation of a user interaction unit (1). In case that the user’s intention can not be fully recognized by the verbal unit, the system will consult the visual perceptual module via OAS in the non-verbal unit (2) and fuse these multi-modal information on the user conversation layer of the interaction unit. Once the user intention is fully recognized, the system creates an interaction unit for itself and tries to provide acceptance: the system first formulates and sends commands to the robot control system or the knowledge base on the domain layer (3) and then generates verbal and non-verbal output on the conversation layer (4, 5) after receiving the execution results. Currently, we have implemented the visualization of facial expressions as the only non-verbal output. Thus, the integration of speech and visual information is mainly performed on the conversation layer of the interaction unit.

In the whole language perception and production process problems may occur, e.g., the semantic meaning of the user interaction unit cannot be resolved or the desired task cannot be executed by the robot system. These problems cannot be handled in a single interaction unit, new interaction units are necessary. Now the question arises as to how to organize the individual interaction units.

4.2 The grounding mechanism

The interaction units have to be organized in a dynamic way since every new contribution that is added to the dialog changes the dialog state. Our dialog model is inspired by the common ground theory of Clark [5]. According to this theory a dialog is carried out in the way that one participant presents an account (presentation) and the other issues the evidence of understanding of the account (acceptance). The grounding process is complete and both dialog participants can go on with a new account only if the acceptance is available. Dialog systems that implement this psychological model ([15], [4]) differ in their way of defining grounding units (the units of the discourse where the grounding takes place) and the organization of these units. We take exchanges in the style of adjacency pairs [14] as the grounding units. These exchanges consist of two interaction units that are initiated by the two dialog partners, respectively. The first interaction unit is the presentation and the second one is the acceptance, e.g., the first interaction unit represents a question and the second one the answer. To organize them we introduce four grounding relations between exchanges: (1) *default*: introducing a new task. A grounded default exchange has no further effect on the grounding of its preceding exchange. (2) *support*: clarifying in case of an ungroundable account. After a support exchange is grounded its initiator will try to ground the preceding one again that is updated with the new information. For example, clarification question in case of an incorrect speech recognition result. (3) *correct*: correct the previous account. As support, if such an exchange is grounded the initiator will try to ground the updated preceding one again. (4) *delete*: delete the previous account. If such an exchange is grounded, all the previous ungrounded exchanges can be deleted.

Each contribution is analyzed in terms of whether it is a presentation or an acceptance. If it is a presentation, then we also need to find out its grounding relation to the preceding exchanges. A presentation initiates the creation of an exchange that is put onto the top of a stack while an acceptance completes the top exchange of the stack. When the top exchange is complete it is popped. Additionally, actions like updating the preceding exchange can be triggered according to these relations. As long as there is an incomplete exchange on the top of the stack, the conversant other than the initiator of the exchange’s presentation will try to ground it. The implemented dialog system enables us to handle clarification questions (as an exchange with support relation to its preceding exchange) and take initiative that is motivated by the robot control system. For example, in case of technical problems of the robot control system the implemented dialog system initiates an interaction unit to report this problem to the user. It does this by encoding the error message into its domain layer and generating output to the user on the conversation layer.

4.3 Resolving object references

In the following we detail how the Dialog System and the OAS cooperate to resolve object references in the user’s utterances.

According to [9] there are three types of informational relations between gesture and speech: *reinforcement*, *disambiguation* and *adding information*. In our work, we focus on the “adding information” relation. When people use gestures to complete the meaning of their utterances they

mostly indicate this intention in the utterance. For example, if a user says “This is my green mug” while pointing at a mug, the word “this” serves as a cue for the listener that he/she is using a gesture to specify the concrete location of the mug. But in case of the subsequent utterance “The mug is my favorite one” the listener usually does not expect a gesture but will search mentally in the dialog history which cup might be meant. According to these two different cases the Dialog System activates either the OAS or the Scene Model to resolve the object reference. This process can be illustrated as a UML activity diagram (see Fig. 3).

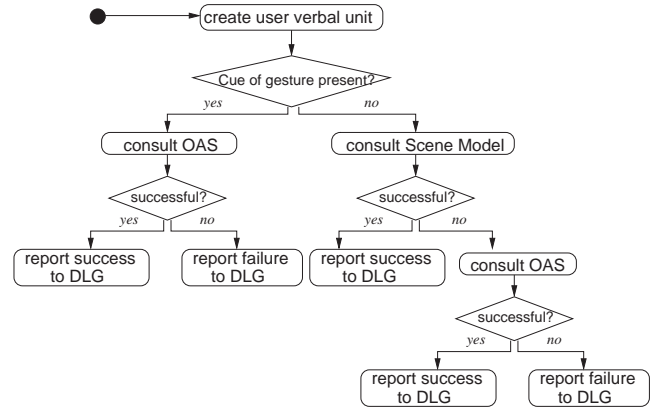


Figure 3: Resolving object references in user’s verbal input (OAS: Object Attention System, DLG: Dialog System)

If there are any cues of the involvement of a gesture in the user’s verbal input, e.g., the word “this” in the example above, this will be explicitly pointed out by the Speech Understanding System. The Dialog System interprets this hint as evidence that the user’s verbal unit needs to consult the non-verbal unit. In the non-verbal unit the Dialog System activates the OAS by sending it the request to resolve the object reference “my green mug”. The OAS activates the gesture recognition module. A successful gesture recognition result helps the OAS to orient the camera towards the position of the user’s hand which enables the OAS to confine the Region Of Interest for its search of a green mug. This search is carried out, as the case may be, either by an appearance-based object recognizer or with the help of salient object features as described in section 5. Subsequently, the OAS sends the search result back to the Dialog System. In case of a positive result the OAS also updates the Scene Model with both symbolic and visual information about the new object “green mug”. According to this result the Dialog System creates a system interaction unit (cf. Fig. 2) to provide acceptance for the user’s input by either acknowledging or by initiating verbal repair.

If there is no evidence of the involvement of a gesture in the user’s verbal input but only some objects to be identified (such as the “mug” in the example “The mug is my favorite one”) the Dialog System will first try to find a corresponding entry in the Scene Model. The query is constructed with all the features of the object present in the current verbal input; in this case, the owner of the cup and his/her relation to this object (favorite). If the object can be found in the Scene Model the Dialog System finishes its processing on the conversation layer of the user interaction unit and creates an

acceptance to the user’s input; if this object is not registered in the Scene Model, the Dialog System activates the OAS to find it in the current scene. This process is described in detail in the following section.

5. OBJECT ATTENTION SYSTEM

In order for a system to be able to acquire knowledge about objects in its environment it needs a mechanism to focus its attention on those objects that the user is currently talking about. In our context we define attention as the ability to select and concentrate on a specific stimulus out of all stimuli that are provided by the environment while suppressing others. The Object Attention System (OAS) therefore needs to coordinate the visual processing results (which currently consist of deictic gestures, object recognition results, and visual object features) and making them accessible for the Dialog System by storing them in the multi-modal Scene Model.

The OAS is activated when the user is verbally referring to an object and the Speech Understanding System has determined that either a gesture is expected or that the robot has to interact with an object autonomously. In order to acquire visual information about objects, BIRON uses an active camera with a maximal opening angle of view of about 50 degrees horizontal and 38 degrees vertical which facilitates only a limited field of view. It can therefore be necessary to re-orient the camera to relevant parts of the current scene which the user refers to. Once the robot has focused its attention on such a so-called *Region Of Interest*, the acquisition of information about this object, like position or view, can be completed. The Region Of Interest is that part of the image that has been specified by a gesture and contains the distinctive feature verbally specified by the user. Our assumption is that it contains an image of the object if the object is known to the robot. If it is unknown the Region Of Interest encloses the verbally specified visual feature (e.g. the “round thing” or a color).

The collected object data then has to be added to the robot’s knowledge base which must allow retrieving stored object information and updating already stored data. Also, additional information given verbally by the user has to be stored. As this knowledge base, subsequently named as *Scene Model*, is crucial for the interaction between the Dialog System and the OAS because it represents BIRON’s long-term memory, it is described in detail in section 6.

In addition to the maintenance of the Scene Model for memorizing tasks, the OAS needs to take care of the coordination of verbal information, gestures, and salient object features (e.g., color, shape, etc.) perceived by the camera, as well as the control of the hardware components like the camera during the object attention phases. This is realized by a *Finite State Machine* (see Fig. 5) where the input is mainly provided by the camera, the Dialog System (cf. section 4), the Speech Understanding System, and the gesture recognition component [7].

A crucial distinction that has to be made during the processing of multi-modal information is that between objects that are already known to the robot and those that are not (see Fig. 4). This is because in the latter case the OAS will have to establish (or ‘learn’) a first link between the verbal symbols describing the object and the percepts while in the former case the object needs to be retrieved from the database.

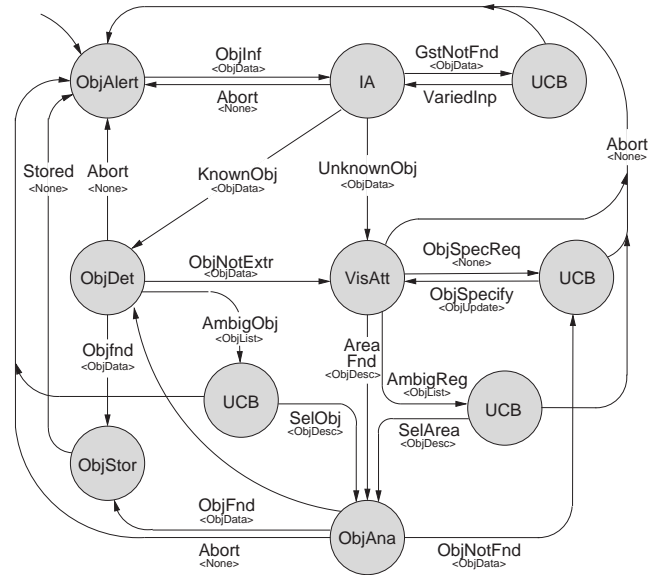


Figure 5: The Finite State Machine of the Object Attention System.

In both cases the OAS is activated on demand by the Dialog System if a gesture is expected or an access to the Scene Model by the Dialog System has failed after. At this moment, the Finite State Machine (see Fig. 5) will be in the idle state *Object Alertness* (ObjAlert). Once the OAS is provided with data by the Dialog System, the Finite State Machine changes to the *Input Analysis* (IA) state. Now, the gesture recognition component is activated and provides the OAS with the user’s hand coordinates and the direction of the corresponding pointing gesture. Thus, an area within the camera image is selected as the Region Of Interest. In case the Dialog System sends a description of the object (e.g., type, color, owner, etc.) to the OAS, a query to the Scene Model is initiated in order to check whether the object type is already known. In the following, we will describe in detail the processing for the case when the object type is known to the robot. The more complex process for the case of unknown objects will be exemplified subsequently.

5.1 Previously known objects

Suppose the user specifies an object type that the system has already stored in its Scene Model. In this case the Scene Model will return all object entries that match the symbolic description of the specified object. In order to verify if one of the returned objects is indeed the object the user refers to, the OAS will need to search for the object in the real scene and compare it with the stored image pattern. This search involves an object detection process for which we are currently using a simple appearance-based object recognizer that is only suitable for a very limited object scenario. It is based on the fast Normalized Cross-Correlation (NCC) algorithm described in [11] which is a simple but fast algorithm that is sufficient for our task at hand. However, in order for the system to work reliably in a more unstructured environment, as for example a real *home-tour* scenario a more sophisticated object recognizer will be needed.

After all appropriate image patterns have been retrieved for the known object type, the Finite State Machine switches

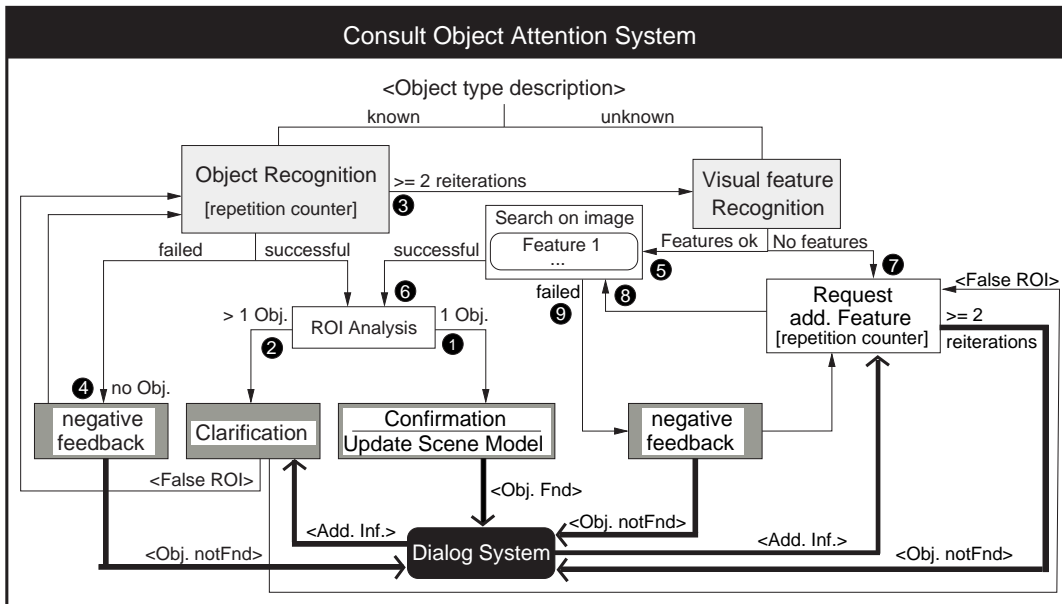


Figure 4: Schematic description of distinction between known and unknown objects

from the Input Analysis state to the *Object Detection* (ObjDet) state. Within the Object Detection state the OAS uses the retrieved image patterns (e.g., for cups) in order to feed them to the object recognizer. At the same time the camera is re-oriented based on the hand coordinates and the pointing direction that are provided by the gesture recognition component. Also, the position of the hand is used to determine the relevant Region Of Interest. Based on this information, the object detection process is initiated by scanning the Region Of Interest for patterns similar to those provided by the database. If an object is found by this procedure a confirmation message is sent to the Dialog System (cf. Fig. 4 (1)) and the Finite State Machine switches to the *Object Store* (ObjStor) state. In this state the position of the object in the scene is updated in the Scene Model. Finally, the Finite State Machine returns to the Object Alertness state and the OAS awaits new orders.

If two or more objects of the same type are found in the real scene during the detection phase (cf. Fig. 4 (2)) in the Object Detection state, the Finite State Machine will switch to the *User callback* (UCB) state. This means, that a message is sent to the Dialog System to clarify which of the found objects was meant by the user. After the Dialog System has provided a more detailed description, the Finite State Machine switches to the *Object Analysis* (ObjAna) state. In this state a new Region Of Interest is determined based on the information from the gesture detection and the extended verbal information. Now the Finite State Machine returns to the Object Detection state and initiates a new search. This cycle is performed until the object is found, or the user aborts the action within the User callback state but at most two times. Then, the Finite State Machine switches to the *Visual Attention* (VisAtt) state (cf. Fig. 4 (3)). If no object is found in the Object Detection state (cf. Fig. 4 (4)) this means that the user is referring to an unknown object which is supposedly similar to the description of objects previously retrieved from the Scene Model. However, since the

object is unknown the Finite State Machine switches to the Visual Attention state, that is also used for the localization of unknown objects if two reiterations have been reached (cf. Fig. 4 (5)).

5.2 Unknown objects

If no object detection is possible because no object entry matching the user's specification has been found in the Scene Model the OAS will search for salient features in the camera image such as colors or shapes by applying different filters that detect salient visual object features as specified by the user. We call these filters *attention maps* following the terminology of [8] where a similar technique is used. The use of these attention maps is coordinated within the Visual Attention state and can help to select Regions Of Interest. The appropriate attention map is selected based on the verbal information (e.g., the color) given by the user (cf. Fig. 4 (6)).

Once a region matching the search criteria (i.e., color) is found within the Region of Interest by the attention map it is selected within a bounding box (cf. Fig. 4 (6)). This bounding box is supposed to contain a view of the retrieved object (e.g., a blue cup) and is stored in the Scene Model (cf. Fig. 4 (1)). Additionally, a confirmation message is sent to the Dialog System.

If the verbal information given by the user is insufficient to determine a Region Of Interest, that is if no visual descriptions are given that can be found by the attention map, (cf. Fig. 4 (7)), the Finite State Machine changes to the User callback state. In the User callback state the OAS sends a request to the Dialog System in order to get more information (e.g., shape, position, ...) about the object which the user refers to. When the user has given a more specific description which is sent by the Dialog System to the OAS, the Finite State Machine returns to the Visual Attention state (cf. Fig. 4 (8)). The User callback state is also reached if more than one Region Of Interest is found (cf. Fig. 4 (2)). Then, the OAS asks the Dialog System to re-

solve this ambiguity. As soon as the OAS has determined the Region Of Interest, the Finite State Machine switches to the Object Analysis state to acquire the position of the object by means of the hand position of the user. Next, the Finite State Machine switches to the Object Store state and stores the extracted view and the position of the object in the Scene Model (cf. Fig. 4 (①)). Then, the Finite State Machine returns to the Object Alertness state to await new orders.

If no Region of Interest is found during the search for visual object features, the Finite State Machine switches to the User Callback state and returns a negative response to the Dialog System. In parallel, the OAS asks the Dialog System for a more detailed object description and re-initiates a second search on the image (cf. Fig. 4 (②)). If for a second time no Region Of Interest is found, the OAS sends a message to the Dialog System, that the search for the referenced object was not successful and returns into its idle state to await new orders from the Dialog System.

6. REPRESENTATION

Information acquired by the Dialog System and the OAS in the ongoing interaction with a user must be stored in an appropriate way. Because the same information from different modalities require different ways of representation the management of such a multi-modal database is a non-trivial task. Our approach to such a database, that we call *Scene Model*, is based on the concept of an active memory [17] since it uses intrinsic processes which allow not only a simple access to the data but also provides intelligent maintenance functionalities. One of the intrinsic processes for example enables the autonomous removal of obsolete information about objects (e.g., the position of a cup three months ago). This *forget* mechanism is quite essential for our application since the robot’s environment is continuously changing.

Within our work we have extended the functionality of the active memory in order to be able to handle the different modalities by storing the same data in different formats. This information that might seem to be redundant at first glance is necessary because the Scene Model is used as BIRON’s long-term memory and both Dialog System and OAS access the data stored in it. For example, consider the color of an object: the Dialog System may store its value in form of a character string, e.g., “blue”; but after finding it in the current scene the OAS may need to store its color value based on the *Hue Saturation Intensity* (HSI) color model. The same holds true for the position of an object, which the Dialog System would store symbolically as “on the table” while the OAS would store its coordinates. The coordinates describing the position of an object are obviously quite useless for the Dialog System when the user asks “where is my blue cup?”. On the other hand, the OAS would not be able to handle the value “blue” when it has to find a blue cup in the camera image.

Consequently, the Scene Model needs to include a component that is able to convert the format of data, the so-called *Modality Converter*. The Modality Converter is a simple yet powerful mechanism that is not only able to convert single object features like the color. It can also search the data base and will return whole object entries matching given descriptions. This may be necessary for example when the OAS fails to detect an object by matching all memorized views

against the current camera image and the object’s color is not yet known to the OAS. Then, in order to extend the search for visual object features, the OAS sends a request for the object’s color to the Dialog System. Subsequently, a search for the newly given color can be performed, after an appropriate conversion of the Dialog System’s response is received by the OAS.

For the conversion process the Modality Converter uses a lookup table (cf. Table 1) that contains for every stored *predicate name* (e.g., color, relation, ...) two attribute fields, in particular a *symbolic description* as well as a *visual feature description*. Since the HSI values might vary for a distinctive verbally named color, the corresponding value field can contain specific values as well as ranges of values. Depending by which module a query is sent the Scene Model returns automatically the adequate description if available. For instance, if the query originates from the Dialog System the Converter will automatically return the attribute “Symbolic”.

Predicate name	Symbolic	Visual
Color	red	330..20,0..1.0,0..1.0 0.0,1.0,1.0
Color	green	135,1.0,0.7
Relation	O_1 under O_2	$O_1.y < O_2.y$

Table 1: Lookup table of the Modality Converter

Note that this converter is a very powerful tool since it does not only convert pre-defined symbol-value pairs but it is also able to learn new associations. In sum, three different responses to queries are possible. ① The converter finds an entry where all data fields match the attributes of the specified object. In this case, it will return the value suitable for the inquiring component. ② The converter finds no valid entry because there is no correspondence of the entry in the other modality as for example for the symbolic name “transparent” which does not have an HSI equivalent. ③ The converter finds no valid entry because it is not yet complete. This usually occurs when after a search for visual object features a new HSI value is stored in the Scene Model for which no symbolic name is yet known. Then, the Dialog System will ask the user for the color name of the Region Of Interest.

7. RESULTS

In order to illustrate our results we present a dialog example where the resolution of object references is involved. In this example, the user asks the robot to pay attention to a mug. Fig. 6 illustrates the dialog flow, the operations of the modules underlying the robot output and the content of the Scene Model in the first, second and third column respectively. We assume that the robot has already recognized the user as Thomas.

In the utterance U1 the word “this” indicates a possible accompanying gesture which can help to specify its meaning. The Dialog System therefore sends a request to the OAS to search for the object mug (①). Upon this request the OAS will first query the Scene Model for an object of the type mug in order to provide a template to the object recognizer (②). In our example, no such object is stored in the Scene Model (③). The OAS now switches to its second searching strategy: search with salient features of the object in the current scene. But since neither salient fea-

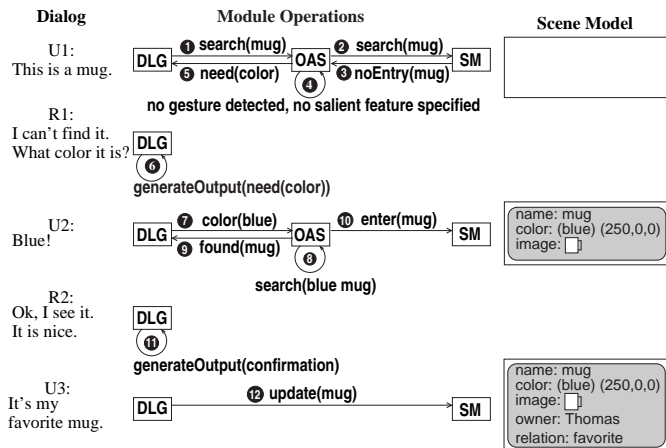


Figure 6: A dialog example (U: User; R: Robot; DLG: the Dialog System; OAS: the Object Attention System; SM: Scene Model)

tures of the mug were specified, nor a gesture was found in the current scene (4), the OAS informs the Dialog System about its need of a salient feature, namely the color (5). The Dialog System will then generate the output R1 to get the requested information from the user (6). Thomas answers the question and the value of the color is sent from the Dialog System to the OAS (7). With this information the OAS successfully finds the object mug (8) and informs the Dialog System about this result (9). At the same time the multi-modal information about the mug is entered in the Scene Model by the OAS (10). The Dialog System can now generate a confirmation (12) as feedback to the user. In U3 the user specifies two further features of the mug, the owner (“mine”) and a description (“my favorite”). This information is directly entered into the Scene Model by the Dialog System since there is no evidence of the involvement of a user gesture, which would indicate that a new object is being specified, and there is only one entry in the Scene Model that matches the described object.

8. CONCLUSION

In order for a mobile robot to be able to communicate in a human-style it needs processing and representation strategies that can deal with multi-modal information. We therefore integrated a Dialog System using multi-modal interaction units with an Object Attention System that is able to resolve object references. The interaction between these modules is based on a multi-modal representation, the Scene Model, which stores the acquired scene information and provides a Modality Converter that not only converts information from one modality to another but also can learn associations between data from different modalities such as color names and HSI values. This powerful mechanism allows on the one hand to use a pre-defined knowledge base while it is on the other hand capable of adapting to new environments by learning new objects and salient features.

The current system still has some limitations. Firstly, the robot cannot yet learn new words, this means, the robot can

only learn objects with known symbolic names. A solution of this problem can be adding a mechanism into the DLG that can store and reuse new words once they are spelled by the user. Secondly, a global 3D coordinate system representing the absolute position of an object in relation to the room is not yet integrated into the OAS. The consequence is, the robot can only find the object again if it is in the position as it learned the object. Thirdly, no navigation system is implemented for the robot so that the robot cannot autonomously move from one location to another to find an object. These limitations are also the motivation for our future work.

9. REFERENCES

- [1] K. Aoyama and H. Shimomura. Real world speech interaction with a humanoid robot on a layered robot behavior control architecture. In *IEEE Int. Conf. Robotics & Automation*, pages 3825–3830, 2005.
- [2] C. Breazeal, A. Brooks, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, A. Lockerd, and D. Mulanda. Humanoid robots as cooperative partners for people. *Int. Journal of Humanoid Robots*, 2004.
- [3] J. S. Brunder. *Towards a Theory of Instruction*. Norton, New York, 1966.
- [4] J. E. Cahn and S. E. A psychological model of grounding and repair in dialog. In *Proc. Fall 1999 AAAI Symposium on Psychological Models of Communication in Collaborative Systems*, 1999.
- [5] H. H. Clark, editor. *Arenas of Language Use*. University of Chicago Press, 1992.
- [6] S. S. Ghidary, Y. Nakata, H. Saito, M. Hattori, and T. Takamori. Multi-modal interaction of human and home robot in the context of room map generation. *Autonomous Robots*, pages 169–184, 2002.
- [7] A. Haasch, N. Hofemann, J. Fritsch, and G. Sagerer. A multi-modal object attention system for a mobile robot. In *Proc. Int. Conf. on Intelligent Robots and Systems*, 2005.
- [8] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *PAMI*, 20(11):1254–1259, 1998.
- [9] J. M. Iverson, O. Capirci, E. Longobardi, and M. C. Caselli. Gesturing in mother-child interactions. *Cognitive Development*, 14(1):57–75, 1999.
- [10] M. Kleinhagenbrock, J. Fritsch, and G. Sagerer. Supporting advanced interaction capabilities on a mobile robot with a flexible control system. In *Proc. Int. Conf. on Intelligent Robots and Systems*, pages 3649–3655, 2004.
- [11] J. P. Lewis. Fast template matching. In *Proc. Conf. on Vision Interface*, pages 120–123, 1995.
- [12] D. McNeill. *Hand and Mind: What Gesture Reveal about Thought*. University of Chicago Press, 1992.
- [13] P. Prodanov and A. Drygajlo. Decision networks for repair strategies in speech-based interaction with mobile tour-guide robots. In *IEEE Intern. Conf. Robotics & Automation*, pages 3052–3057, 2005.
- [14] E. Schegloff and H. Sacks. Opening up closings. *Semiotica*, pages 289–327, 1973.
- [15] D. Traum. *A Computational Theory of Grounding in Natural Language Conversation*. PhD thesis, University of Rochester, 1994.
- [16] D. Traum and P. Dillenbourg. Miscommunication in multi-modal collaboration. In *Proc. AAAI Workshop on Detecting, Repairing, And Preventing Human-Machine Miscommunication*, 1996.
- [17] S. Wrede, M. Hanheide, C. Bauckhage, and G. Sagerer. An Active Memory as a Model for Information Fusion. In *Proc. Int. Conf. on Information Fusion*, pages 198–205, 2004.

The Influence of Subjects' Personality Traits on Personal Spatial Zones in a Human-Robot Interaction Experiment^{*}

Michael L Walters, Kerstin Dautenhahn, René te Boekhorst, Kheng Lee Koay, Christina Kaouri, Sarah Woods, Christopher Nehaniv, David Lee, Iain Werry

*Adaptive Systems Research Group
University of Hertfordshire
Hatfield, Herts. UK*

{M.L.Walters, K.Dautenhahn, R.TeBoekhorst}@herts.ac.uk

Abstract – In the present study we investigated human-robot and robot-human approach distances. We found that subjects' personality profiles influence personal spatial zones in human-robot interaction experiments. We tested two hypotheses: First, we predicted that approach distances preferred by humans when interacting with a robot would be comparable to those preferred when humans interact socially with each other. Our experiments involving humans interacting with a mobile robot confirm this hypothesis. However, surprisingly, a large minority of subjects in the experiments took up positions which were significantly closer, suggesting that they were not treating the robot as a 'social entity'. We then tested the hypothesis that common personality factors exist which could be used to predict subjects' likely approach distance preferences. The subjects' personalities were assessed using several traits from the three-factor Eysenck personality model. Further analysis of the data identified four new factors, different from Eysenck's model, tentatively labeled "Proactiveness", "Social Reluctance", "Timidity" and "Nervousness". When testing for correlations between approach distances and personality data, "Proactiveness" correlates with social distance, i.e. subjects that score higher on this factor come less close to the robot. We discuss the potential suitability of personality factors to predict approach distances in human-robot interaction experiments.

Index Terms - Human-Robot Interaction, Social Robot, Social Spaces, Personal Spaces

I. INTRODUCTION

Studying social and personal spaces with regard to robots, designed for use in the home, is a particular area of research within the wider field of Human - Robot Interaction (HRI). In the near future, it is anticipated that robots will increasingly be used for applications in office and domestic environments. Therefore they will be required to work alongside and interact closely with the human residents [1]. As the study of socially interactive robots is relatively new, there is not a large body of established theories, methods and research experience to draw upon, so experimenters in the field usually use existing research into human-human social interactions as a starting point. These methods and results, along with later research, have provided a guide for more recent research, studies and investigations into human reactions to and attitudes towards robots [2] - [16].

A. Human-Robot Social Spaces

The main emphasis of our research is on the physical, spatial, visual and audible *non-verbal* social aspects of robots interacting socially with humans. In particular, we are interested in studying human-robot social spaces and distances. Hall [17] described a basis for research into social and personal spaces between humans, and later work in psychology has demonstrated that social spaces substantially reflect and influence social relationships and attitudes of people. Embodied non-verbal interactions, such as approach, touch, and avoidance behaviours, are fundamental to regulating human-human social interactions [18]. Spatial zones among people are strongly influenced by cultural factors. The generally recognized *personal spatial zones* between humans are well known and are summarized (for northern Europeans) in Table 1 from Lambert [19].

TABLE 1
HUMAN-HUMAN PERSONAL SPATIAL ZONES

Personal Spatial Zone	Range	Situation
Close Intimate	0 to 0.15m	Lover or close friend touching
Intimate Zone	0.15m to 0.45m	Lover or close friend only
Personal Zone	0.45m to 1.2m	Conversation between friends
Social Zone	1.2m to 3.6m	Conversation to non-friends
Public Zone	3.6m +	Public speech making

This paper presents our exploratory research into human-robot social spaces, investigating whether human-human personal spatial zones transfer to human-robot interaction. As a starting point we have compared human-robot approach distances to those that would be expected for the case of a human approaching another human. A working hypothesis that human-robot interpersonal distances would be comparable to those found for human-human interpersonal distances was used; cf. Walters et al. [19], Christensen and Pacchierotti [21]. We expected that in scenarios designed for direct human-robot interaction, people would assume distances that on average correspond to the 'Social' or 'Personal' zone (similar to the distances people use when having face-to-face conversations with each other) thus treating the robot as a 'social being' with respect to social distances.

II. HUMAN-ROBOT SOCIAL DISTANCE EXPERIMENTS

^{*} The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

The human-robot social space experiments were performed prior to a separate series of experimental sessions studying human-robot interactions in a range of task based scenarios. The subject sample set consisted of 28 adult volunteers (staff members and students from University of Hertfordshire). The sample was balanced for age, gender and whether subjects had a technology related background. All subjects completed consent forms and were not paid for participation.

A. The Robot:

The robot used for this study was a commercially available PeopleBot™ robot which is mechanistic in appearance (see Fig. 1). This is a human scaled robot, 1.1m tall, with a camera with pan and tilt capabilities. The robot was also fitted with three banks of eight sonar range finders which allow the robot to sense objects at low level (approximately 0.25m from the ground) all around, and at high level (at a height of approximately 1m) in front. The sonar sensors are particularly good at sensing soft targets, such as humans and semi hard materials such as walls, furniture etc, and are primarily used for object avoidance and safe movement in environments containing humans. The robot is steered by two differential driving wheels, and has a caster wheel at the back and front to provide stability. The only anthropomorphic feature of the robot was a lifting arm, with a hook type end-effector, to allow the robot to fetch and carry small objects in specially adapted pallets. The robot was operated under remote control by two hidden operators. This is commonly called Wizard of Oz (WoZ) and is a technique that is widely used in HRI studies. It provides a very flexible way to implement complex robot behaviour within a quick time-scale [22][23]. The main advantage is that it saves considerable time over programming a robot to carry out complex interactions fully autonomously.

At the start of each experiment the robot was driven to the same fixed position in the room for each approach distance test. This was achieved by using the table in the corner of the room (position 5 in Figure 3) as a stop position reference for the robot's sonar range sensors. The robot could then therefore be driven towards the corner, until it stopped at a fixed distance from the corner.

B. Experimental Method:

The experimental sessions took place in a conference room at the University premises, which was converted and furnished to resemble a domestic sitting room as far as was possible. One end of the room was partitioned off using shelf units, cupboards and high screens to form a control area for the robot operators. Marks were made on the floor using masking tape along the diagonal of the experiment room, and scale marks made at 0.5m intervals between them (Figs. 2 and 3). The experiments were supervised by an experimenter who introduced and explained the tests to be carried out to the subject. Otherwise, she interfered as little as possible with the actual experiment.

The human-robot comfort and approach distances were estimated from video records of the sessions, rather than having the experimenter making intrusive measurements or notes during the experimental sessions.



Fig 1. The PeopleBot™ robot used in the experiments.



Fig 2. View of the simulated living room showing robot and the 0.5m scale marked diagonally on the floor

Each experimental session followed the same format: 1) Entry to room and introduction of robot, 2) Co-habitation and initial questionnaires; While the subject was filling in the first questionnaires, the robot wandered randomly around the test area to acclimatise the subject to the robot, for a period of five to ten minutes prior to the distance tests, 3) Comfort and social distance tests, 4) Various other HRI task scenarios and questionnaires¹.

For measuring the human subject's comfort threshold distance when approaching the robot, the robot was driven to point 5 (Figure 3), next to the corner table and turned to face along the distance scale towards point 4 (Figure 3). The subject was told to start at point 4 and to move towards the robot until he or she felt that they were at a comfortable distance away from the robot (Figure 4). The instruction used was "Move towards the robot as far as you feel comfortable to do so". Next, they were told to move as close to the robot as they physically could (if not already in that position); "Move as close to the robot as you physically can". Then they were told to move away again to a comfortable distance; "move back to your most comfortable distance". They were then told

¹ These latter parts were carried out for separate HRI investigations and are therefore not considered in this paper.

to repeat these steps once again as a consistency check. The comfortable approach, closest physical and comfortable withdrawal distances were measured for each of the two tests by later close observation of the video records. The distances were estimated to the nearest 0.25m (accuracy $\pm 0.125m$).

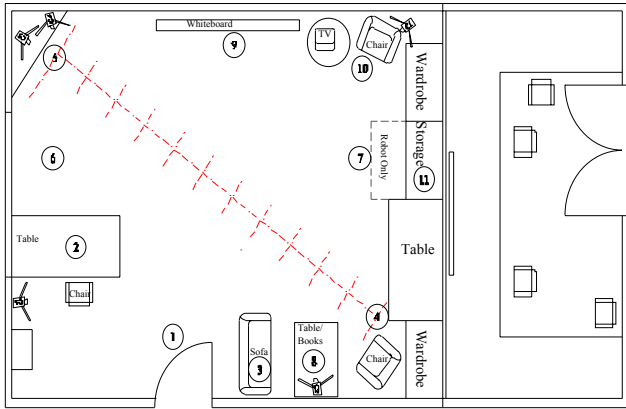


Fig 3. Plan view of simulated living room layout. Comfort distance tests carried were out along the marked diagonal line.

For the human-robot approach distance experiments two measurements for the comfortable approach distance and two for the comfortable withdrawal distance were obtained. In practically all cases subject’s withdrawal distances were within 0.25m of their comfortable approach distances. The four approach and withdrawal distances were then aggregated to produce a single mean comfortable approach distance value for each subject.



Fig 4. Human-robot approach distance experiment; a human subject approaching the robot

A second set of comfortable approach distance measurements were then made for the situation where the robot approached a stationary human subject. The subject was asked to stand at position 4), and the robot was driven to position 5) (the diagonally opposite corner of the room. see Fig. 3). The subject was then asked to say “Stop” when the robot came as close as they felt was comfortable. The robot was then driven directly towards the subject at a speed of approximately 1 meter per second. When the subject said “Stop” the robot was stopped as quickly as the WoZ operators

could react. This usually involved an overshoot of 0.5m or more. The distance of the robot from the subject at the instant when the subject actually said “Stop” was estimated from the video records of the experiment. By using the video system stop frame facility and the 0.5m scale marks on the floor it was possible to estimate the distance to the nearest 0.25m (accuracy $\pm 0.125m$). The robot-human approach distance experiment was also repeated twice as a consistency check. Two robot-human comfortable approach distance measurements were obtained, which were then aggregated to obtain a single mean distance value for each subject.

C. Results

The means of the four human to robot comfortable approach distance results obtained were calculated for each subject and a frequency histogram was plotted, with the ranges set at 0.25m intervals (consistent with the accuracy of the measurements). The results are shown in Figures 5 and 6. For the case of the robot approaching the human the means of the two distances for each subject are shown in Figure 7. There was no robot to human approach distance less than 0.5m as the robot’s anti-collision safety system prevented it moving closer than 0.5m to a human (or any other object).

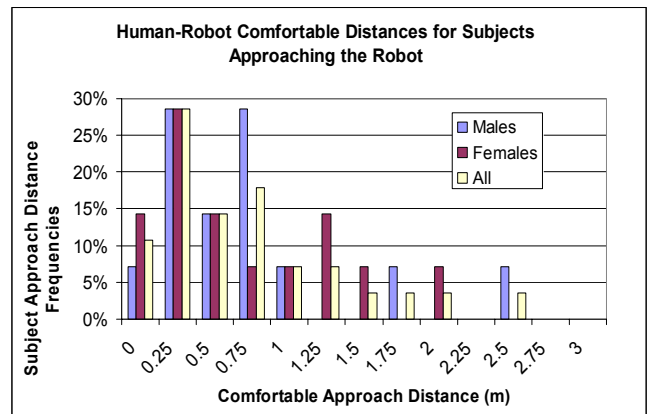


Fig 5. Comfortable distance frequencies for subjects approaching the robot. (Shown as percentages of the subject sample set: N= 28, M= 14, F= 14)

The approach distance to the robot for the majority of subjects (60% total) was within the expected ranges for comparable human-human social distances, corresponding to either the personal or social spatial zones. However, approximately 40% of subjects approached the robot to a distance of less than 0.45m. Also, 38% of the subjects allowed the robot to approach right up to the 0.5m limit set by the robot’s safety system. The fact that they did not stop the robot from physically approaching so closely indicates that the robot did not make these subjects feel threatened or uncomfortable. Indeed, if another, unfamiliar human (a stranger) was to approach to the same close distance; most humans would start to feel distinctly uncomfortable and threatened. Practically all the subjects stated that they did not feel threatened by the robot (and only a minority wished to become intimate with the robot in the sense of having the robot as a friend or companion). It is probable therefore, that this large minority of

subjects did not relate to the PeopleBot™ in terms of the normal social distances between humans, i.e. reflecting a conversation between friends or acquaintances.

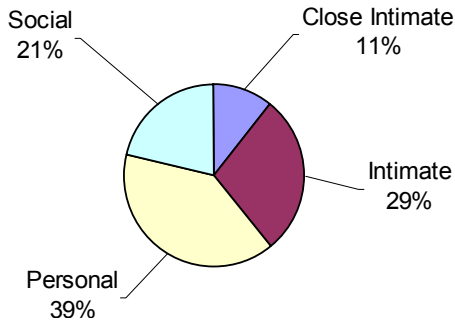


Fig 6. Human to robot comfortable approach distances categorised into Hall's personal spatial zones. (Shown as percentages of the subject sample set: N = 28)

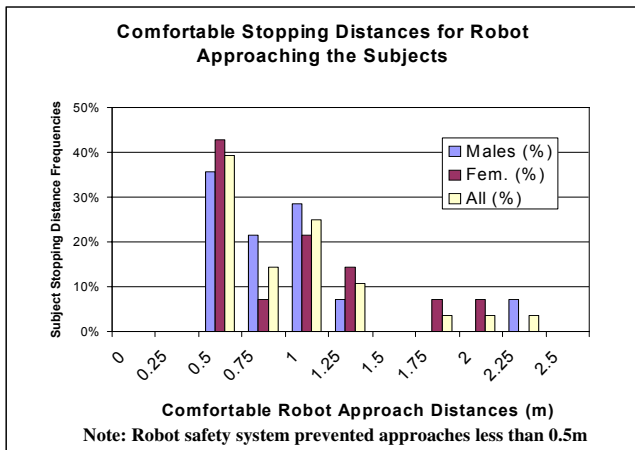


Fig 7. Social (stopping) distance frequencies when the robot approaches the human subjects. (Shown as percentages of the subject sample set: N=28, M=14, F=14).

III. SOCIAL DISTANCE AND SUBJECTS' PERSONALITIES

In order to address the issue of personality, we chose Eysenck's Three-Factor Psychoticism, Extroversion and Neuroticism (PEN) model as a starting point [24]. In Eysenck's view, personality types are not categories that a few people fit; rather, types are dimensions that span a space in which persons can be pinpointed at all possible positions [25]. Types tend to be normally distributed, meaning that they can take continuous values and most people fall around the average mark. From his studies with human subjects Eysenck concluded that personality can be understood in terms of three basic personality factors, which in turn are composed of a number of traits. Due to the time constraints for the HRI experiments not all 27 Eysenck traits were used. Also, since the subjects were rating themselves, certain traits were not considered suitable for self-assessment (e.g. antisocial). Moreover, we considered only traits that could be applied to rating human as well as robot personality [16]. The following traits were chosen that seemed most relevant to the present study:

Psychoticism: The loss of distortion of reality and the inability to distinguish between reality and fantasy. This is not a dimension like the other two (it does not consist of polar opposites) - it is present in all individuals to some degree. The traits associated are: Aggressive, Cold, Egocentric, Impersonal, Impulsive, Antisocial, Un-empathetic, Creative and Tough-minded. The traits selected to be used for our study were: Aggressiveness, Impulsiveness and Creativity.

Extroversion: Degree to which a person is outgoing and participative in relating to others. Traits associated comprise: Sociable, Lively, Active, Assertive, Sensation Seeking, Carefree, Dominant, Surgent and Venturesome. The traits selected to be used for our study were: Sociability, General Activity Level, Assertiveness, Excitement-Seeking and Dominance.

Neuroticism: An individual's adjustment to environment and stability of behaviour over time. The traits associated are: Anxious, Depressed, Guilt Feelings, Low Self Esteem, Tense, Irrational, Shy, Moody and Emotional. The traits selected to be used for our study were: Anxiety, Tension, Shyness and Emotional Vulnerability.

The subject personality questionnaire required the participants to rate themselves in terms of the 12 different personality traits using a 5-point Likert scale. Subjects were all informed that this information would be treated confidentially and would not be linked to their real name during any stage of the evaluation.

The score for each personality factor (F) for every individual subject, was determined by adding up the score for each of the selected (Eysenck) traits (T) for that particular factor, and dividing by the number of selected traits (N) involved (Formula 1):

$$F = \frac{1}{N} \sum_{n=1}^N T_n \quad (1)$$

The three factors thus produced for each subject were then each rounded to the closest 5-point Likert scale (integer) values to create an individual personality vector for each subject (Eysenck & Eysenck, 1985, p. 192).

A. Results of Personality Questionnaires

Note, instead of the 27 traits used by Eysenck, only 12 of these were measured in the present study. This means that the combined traits used by us may not fully reflect the original Eysenck factors. To check this, we performed a confirmatory factor analysis in which it was assessed in how far the correlation structure of the measured variables fitted with the original factor model of Eysenck. As suspected, none of the chosen goodness-of-fit indices ((Adjusted) population Gamma = (0.664), 0.768, Joreskog GFI = 0.611, Joreskog AGFI = 0.438, Bentler-Bonnet Normed Fit Index = 0.224, BB Non-NFI = 0.083, BB Comparative FI = 0.252, RMSEA = 0.123) lend support for the model and both the ML- and independence model Chi square were highly significant (resp. 127.27 and 163.91, with degrees of freedom of 54 and 66). We therefore decided not to base our interpretations on Eysenck's model, but to analyse the correlations among the 12 selected variables on their own right.

TABLE 2
LOADINGS OF FACTORS ABSTRACTED FROM THE 12 MEASURED TRAITS.

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Sociability	0.366230	-0.728960	-0.175871	-0.026398
Shyness	-0.658963	0.542248	0.016767	0.176433
Vulnerability	-0.424035	-0.559047	-0.052254	-0.472986
General Activity Level	0.677153	0.223206	0.251580	0.118710
Assertiveness	0.444187	0.060199	-0.604208	0.405537
Anxiety	-0.249595	-0.575984	-0.086875	0.629583
Tension	-0.596330	0.023251	-0.137943	0.501070
Creativity	0.723497	-0.104735	0.348556	0.060268
Excitement- Seeking	0.652433	0.271774	0.196174	0.044795
Dominance	0.514370	-0.064347	-0.699076	-0.211223
Aggressiveness	0.143104	0.485739	-0.590123	-0.175436
Impulsiveness	0.887287	-0.065932	0.125375	0.204800
Expl.Var	3.843357	1.852954	1.501525	1.204514
Prp.Totl	0.320280	0.154413	0.125127	0.100376

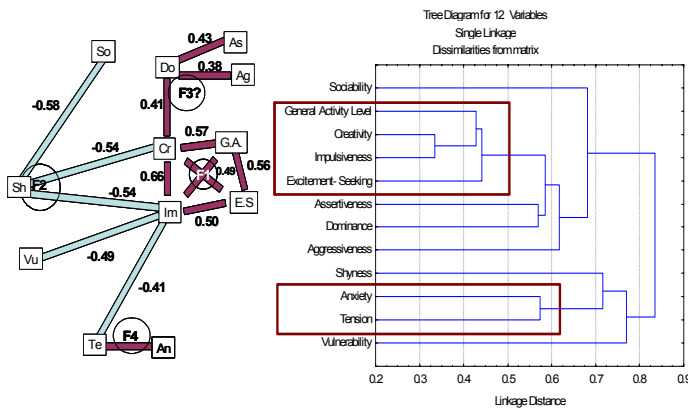


Fig 8. Left: Significant ($p < 0.05$) Spearman Rank Correlations (r_s) among the 12 attributes. Dark bars: positive correlations, light bars: negative correlations. Length of the bars is proportional to the correlation coefficients for $r_s > 0$ and proportional to $1-r_s$ for $r_s < 0$. So = Sociability; Sh = Shyness; Vu = Vulnerability; G.A. = General Activity Level; As = Assertiveness; An = Anxiety; Te = Tension; Cr = Creativity; E.S. = Excitement Seeking; Do = Dominance; Ag = Aggressiveness; Im = Impulsiveness. Note the correspondence with the factors F1 – F4 from the factor analysis (Table II). Right: Cluster analysis based on $1-r_s$ as distance metric and Ward’s Average as cluster criterion. A cluster analysis on the raw data instead of the correlation coefficients gave largely similar results.

An exploratory factor analysis on the 12 traits shows that 70 % of the variance in the data can be explained by four factors (Table 2). The main traits building up the first factor are Creativity and Impulsiveness. At first sight, this seems to correspond with the “Psychoticism” factor. However, instead of Aggressiveness, General Activity Level and Excitement Seeking also contribute strongly to this factor. We tentatively suggest characterizing this combination as “Proactive” attitudes. This is backed up by the fact that “Shyness” correlates negatively with this factor. Factor 2 appears to reflect the degree of what might be called “Social reluctance”: Shyness contributes relatively strongly to it and there is a strong negative correlation with Sociability. Factor 3 seems to characterize “Timidity”, as Assertiveness, Dominance and Aggressiveness are all associated negatively with it.

“Nervousness” typifies factor 4: both Anxiety and Tension load rather high on it.

The factor analysis is backed up by a Principle Component Analysis (PCA; the principle components are almost identical to the factors), which is based on less assumptions and a non-parametric approach (clustering on Spearman Rank correlations; Fig. 8).

When testing for correlations between approach distances and personality data, “Proactive” is the only factor that correlates with social distance, in the sense that subjects that score higher on this factor come less close to the robot ($r_s = 0.647$, $p < 0.05$). Also the effects of gender, age and technical or robotics experience were investigated. Although males appear to score higher on the second (“Social reluctance”) factor (Mann Whitney U Test, $U = 48$, $z = 2.389$, $p < 0.002$), none of these other demographic factors associated significantly with social distance.

IV. CONCLUSIONS

We have found that a majority of human subjects (60%) when approaching a robot, or when being approached by a robot, prefer approach distances that are compatible with those expected for normal social interactions between humans. This partially confirms our original hypothesis in that it seems that humans respect human-robot approach distances in a way which is comparable to human-human social distances. However, in our experiments a large minority of subjects (40%) took up an initial approach distance to the robot, which was so close that it would be perceived as either threatening (if involving strangers) or intimate (in the case of close friends) if observed between two humans. These subjects clearly did not perceive the robot in a way that was comparable with normal human-human social distances, which might imply that these subjects did not perceive the robot as a ‘social entity’ with respect to distances in the same way as another human being.

We studied subjects’ personalities to see if there were common factors, which could be used to predict the likely approach distance preferred by the subjects. Factor analysis resulted in four factors that we tentatively label “Proactiveness”, “Social Reluctance”, “Timidity” and “Nervousness”. Correlations with the social distance experiments show a positive correlation for “Proactive”, i.e. the more proactive a person judged him/herself the longer the human-to-robot approach distances measured.

At this stage our characterization of these four factors, as an alternative to Eysenck’s factors, is preliminary. For this particular study, we do not suggest to use them as a universal scale for human robot interactions. Potentially, factors might be identified as being most suitable for human-robot interaction studies; possibly specific to particular contexts, task environments, particular robots, and/or experimental settings, but this requires deeper analysis and confirmation in future studies (for an example of such a scale, see [27]). Moreover, the sample of subjects we used was self-selected (University staff/students). A subject sample that is more representative of ‘potential users’ of a robot companion might yield different results and also cultural differences might have to be taken into account [28]). Also, in future work we need to

consider that the markings on the floor could influence subjects' judgements. Furthermore, social distances may be affected by the robots' appearance, subjects' own identification of whether they are seen as social entities, tasks of robots and so on. These factors may obviously play a role in human – robot interactions and are important topics for further studies. However, if we succeed in identifying and confirming a set of such factors, then based on a person's personality assessment, one could adjust human-robot distances according to the subject's personality profile. This would provide an important step towards personalized robot companions [26].

REFERENCES

- [1]. T. Fong, I. Nourbakhsh, and K. Dautenhahn, (2003). A survey of socially interactive robots, *Robotics and Autonomous Systems*, vol. 42, pp. 143-166.
- [2]. S. Thrun, (1998). When robots meet people, *IEEE (see also IEEE Expert) Intelligent Systems*, vol. 13, pp. 27 - 29.
- [3]. J. Goetz and S. Kiesler, (2002). Cooperation with a Robotic Assistant. *Proc. CHI'02 Conference on Human Factors in Computing Systems*, New York, USA, 2002.
- [4]. M. Scopelliti, M. V. Giuliani, A. M. D'Amico, and F. Fornara, (2004). "If I had a robot at home.... Peoples' representation of domestic robots", in *Designing a more inclusive world*. S. Keates, J. Clarkson, P. Langdon, and P. Robinson, Eds., Springer, pp. 257-266.
- [5]. K. Severinson-Eklundh, A. Green, H. Hüttenrauch, (2003). Social and Collaborative Aspects of Interaction with a Service Robot. *Robotics and Autonomous Systems* 42 (2003) pp. 223-234.
- [6]. Y. Nakauchi and R. Simmons, (2002). A Social Robot that Stands in Line, *Autonomous Robots*, Vol. 12, No. 3, pp. 313 -324.
- [7]. C. L. Breazeal, (2002). *Designing sociable robots*. Massachusetts: The MIT Press, 2002.
- [8]. B. Friedman, P. H. Kahn (Jr.), and J. Hagman, (2003). Hardware companions? What online AIBO discussion forums reveal about the human-robotic relationship, *Digital Sociability*, vol. 5, pp. 273-280.
- [9]. T. Kanda, H. Ishiguro, M. Imai, and T. Ono, Body Movement Analysis of Human-Robot Interaction, *International Joint Conference on Artificial Intelligence (IJCAI 2003)*, pp.177-182.
- [10]. H. G. Okuno, K. Nakadai, and H. Kitano (2002). "Realizing Audio-Visually Triggered ELIZA-Like Non-verbal Behaviors," *PRICAI2002*, LNAI 2417, Springer-Verlag, pp.552-562.
- [11]. S. Woods, K. Dautenhahn, J. Schulz, (2004). The Design Space of Robots: Investigating Children's Views. *Proc. IEEE Ro-man 2004*, IEEE Press, pp. 47-52.
- [12]. P Hinds, T Roberts, H Jones, (2004). Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task. *Human-Computer Interaction*, Vol. 19 pp151-181.
- [13]. Z. Khan, (1998). Attitude towards intelligent service robots, NADA KTH Technical Report, Stockholm 1998.
- [14]. K. Dautenhahn, (2002). Design Spaces and Niche Spaces of Believable Social Robots. *Proc. IEEE Ro-man 2002*, IEEE Press, pp. 192-197.
- [15]. T Minato, M Shimada, H Ishiguro and S Itakura, (2004). Development of an Android Robot for Studying Human-Robot Interaction. *Proc. IEA/AIE Conf. 2004*, pp. 424-334
- [16]. K. Dautenhahn, K. L. Koay, R. te Boekhorst, M. Walters, (2005). Perception of Robot 'Personalities' Based on Robot's Behaviour Styles. Submitted for publication.
- [17]. E.T. Hall, (1966). *The Hidden Dimension: Man's Use of Space in Public and Private*. The Bodley Head Ltd, London, UK.
- [18]. E. T. Hall, (1968). *Proxemics*. *Current Anthropology* 9(2-3): 83-108. 1968
- [19]. D Lambert, (2004). *Body Language*. HarperCollins.
- [20]. M Walters, K Dautenhahn, K. L. Koay, R. te Boekhorst, C. Nehaniv, I. Werry, D. Lee, (2005). Close Encounters: Spatial distances between People and a Robot of Mechanistic Appearance. Submitted for publication.
- [21]. H. I. Christensen, E Pacchierotti, (2005). Embodied social interaction for robots. *Proc. AISB'05 Symposium on Robot Companions: Hard Problems and Open Challenges in Human-Robot Interaction*, pp.40-45.
- [22]. B. Robins, K. Dautenhahn, J. Dubowski (2004). Investigating Autistic Children's Attitudes Towards Strangers with the Theatrical Robot - A New Experimental Paradigm in Human-Robot Interaction Studies. *Proc. IEEE Ro-man 2004*, IEEE Press, pp. 557-562.
- [23]. A. Green, H. Hüttenrauch, K Severinson Eklundh (2004). Applying the Wizard of Oz Framework to Cooperative Service discovery and Configuration. *Proc. IEEE Ro-man 2004*, IEEE Press, pp. 575-580.
- [24]. H. J. Eysenck, (1991). Dimensions of personality: 16, 5, or 3?—Criteria for a taxonomic paradigm, *Personality and Individual Differences*, 12, 773-790.
- [25]. H. J. Eysenck and M. W. Eysenck, (1985). *Personality and individual differences: A natural science approach*. New York: Plenum, 1985.
- [26]. K. Dautenhahn, (2004). Robots We Like to Live With?! - A Developmental Perspective on a Personalized, Life-Long Robot Companion , *Proc. IEEE Ro-man 2004*, IEEE Press, pp. 17-22.
- [27]. T. Nomura, T. Kanda, T. Suzuki and K. Kato (2004). Psychology in Human-Robot Communication: an attempt through investigation of negative attitudes and anxiety towards robots. In *Proc. 13th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 35-40.
- [28]. C. Bartneck, T. Nomura, T. Kanda, T. Suzuki and K. Kato (2005). A cross-cultural study on attitudes towards robots. *Proc. 11th International Conference on Human-Computer Interaction* (to appear).

Is This Robot Like Me? Links Between Human and Robot Personality Traits¹

Sarah Woods, Kerstin Dautenhahn, Christina Kaouri, René te Boekhorst, Kheng Lee Koay

Adaptive Systems Research Group, School of Computer Science

University of Hertfordshire, United Kingdom

{s.n.woods, k.dautenhahn} @herts.ac.uk

Abstract - A relatively unexplored question for human-robot social interaction is whether a robot's personality should match that of the human user, or be different in the sense that humans do not want the robot to be like them. In this study, 28 adults interacted individually with a non-humanoid robot that demonstrated two robot behaviour styles (Socially Interactive, Socially Ignorant) in a simulated living room situation. Questionnaires assessed the extent to which adult ratings of their own personality traits were similar or different to the two robot behaviours. Results revealed that overall subjects did not view their own personality as similar to either of the two robot behaviour styles. Subjects viewed themselves as having stronger personality characteristics compared to the two robot behaviour styles. Important group differences were found, factors such as subject gender, age and technological experience were important in how subjects viewed their personality as being similar to the robot personality. Design implications for future studies are discussed.

Index Terms: Human-robot interaction, personality traits, social interaction, robot companion, Eysenck model of personality

I. INTRODUCTION

There are mixed opinions as to how humanlike a robot should be designed to engage in believable human-robot interactions [1, 2], and there are many design considerations to be taken into account both in terms of physical appearance and behavioural competencies [3]. The personality of a social robot is an important research domain for robot designers and has received only limited attention to date. There is no universally accepted definition of personality, but it can be broadly defined as a collection of individual differences, dispositions and temperaments that have consistency across situations and time [1, 4]. Research studies have found that people tend to assign personality attributes to computers, agents, and robots which could assist the user in understanding its behaviour by shaping the users' expectations about the interaction experience [3, 5, 6].

Those studies that have considered personality in relation to human-robot interaction have indicated that the personality of a robot should match its design purpose. For example, Goetz and Kiesler [7] found that people enjoyed robot interactions more with a happy robot, but were more likely to follow instructions from a serious robot. Also, correlations have been found between people's personalities and a virtual

agent's social behaviour style [8]. In the same essence that people are uncomfortable with human personality styles that are unpredictable and inconsistent with situational contexts, it would follow that this relationship could be the same between human-robot interactions. To determine whether this relationship exists in human-robot interactions, it could be the case that humans try to match and project their own personality attributes and styles to that of a robot to create an engaging interaction that they feel at ease with and can make sense of. Alternatively, humans may not want to perceive themselves as being similar to a robot in terms of personality attributes. If this is the case they may infer different personality traits, or no personality traits, in the fear of losing their own identity, and wanting to remain unrelated to a robot.

The measurement of personality remains an area of controversy as there are many debates on the number of dimensions that define personality [9-12]. The widely used PEN/3 factor model of personality assessment was chosen for the present study [11]. Eysenck's view of personality was that people do not fit discrete categories, but instead there are dimensions of personality on which all individuals differ. Based on extensive research studies, the PEN model of personality is comprised of personality traits which intercorrelate and make up superfactors called "types", termed extraversion (E), neuroticism (N) and psychoticism (P). The extraversion vs. introversion factor is associated with the degree to which a person is outgoing and participative in relating to others. Neuroticism vs. emotional stability is related to an individual's adjustment to the environment and stability of behaviour over time. Finally, psychoticism is related to the loss of distortion of reality and the inability to distinguish reality and fantasy. Psychoticism is not a dimension like the other two factors, but is said to be present to some degree in all individuals. The three superfactors are independent dimensions. Our analysis in this paper is inspired by the PEN model to explore the relationship between users' personality traits, and their perceptions of robot personality traits, for two behaviour styles.

II RESEARCH QUESTIONS

Research findings suggest that a robot's behaviour should fit the context and its task performance. Further, there is evidence that humans use personality as a social tool for

¹ The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

interpreting and explaining others behaviour. Although research has shown that humans do attribute personality traits to robots, we are not aware of any research that has considered the nature of the assignment of these traits, and the possible relationship between the user and robot. For example, do humans assign their own personality traits to explain a robot's behaviour and intentions or assign contrasting personality traits in an effort to keep their own personality and identity separate to a non-living robot? This could have design implications in terms of matching robot behaviour and interaction style with desirable personality traits from a user perspective. The research questions for this study were:

- 1) Are there significant *differences* between participant personality traits and assigned robot personality traits?
- 2) Is there a *relationship* between human and robot personality?
- 3) If humans do project their own personality onto robots, does this attribution depend on the way the robot behaves?
- 4) What are the design implications for robots based on the findings from personality theory?

To investigate these research questions, an experiment was conducted where adults interacted with a robot in two different contexts (Negotiated Space Task and Assistance Task) for two contrasting robot behaviour styles (Socially Ignorant, Socially Interactive), in a simulated living room scenario. Participants completed questionnaires both pre and post robot interaction to assess their own personality, and the personality of the robot, based on Eysenck's PEN model.

III. METHOD

Design: Single human participants took part in this study in a simulated living room scenario at the University of Hertfordshire during July and August 2004 (see fig. 1). A commercially available, human-scaled PeopleBot™ robot was used to evaluate differences in participants' social behaviour and interaction styles when the robot displayed two contrasting behaviours (Socially Ignorant, Socially Interactive). The simulated living room was supervised by an experimenter at all times. The role of the experimenter was to explain the robot trials to the subject. If the subject initiated an interaction and wanted to ask a question, only then would the experimenter respond. Questionnaires that are relevant to this part of the study were an introductory questionnaire about subject demographics, subject personality questionnaire and the robot personality questionnaire.

Sample characteristics: See Table 1

Experimental Procedure: Introduction - A general welcome phase where the robot was introduced to the subject when they entered the simulated living room. An information sheet was given to the subject to read, along with a consent form to be signed, an Introductory Questionnaire and a Subject Personality Questionnaire to be completed. The robot moved around the room whilst the subject completed these initial questionnaires to familiarise them to the robot.

The Main Trial consisted of two tasks, a Negotiated Space Task and an Assistance Task (see fig. 2), which were repeated with the two contrasting robot behaviour styles (Socially Ignorant and Socially Interactive). The Negotiated Space Task

involved the robot moving in the room (either with a Socially Ignorant or Socially Interactive behaviour style) while the subject went through a pile of books placed on a table, remembering one title at a time, walking over and writing down each title on the whiteboard. The Assistance Task involved the subject sitting at a table copying book titles from a whiteboard onto a piece of paper and underlining specific letters with a red/highlighter pen. The robot was responsible for bringing the missing red/highlighter pen to the table. The two tasks were chosen as they match the two key scenarios studied in the Cogniron project. At the end of these two tasks, the subject completed a robot personality questionnaire. The main trial was then repeated with the alternate robot behaviour style. Trials lasted for approximately one hour.



Fig. 1 Simulated living room for the human-robot interaction trials



Fig 2. The negotiated space (left) and assistance task in the robot trials.

TABLE I: Sample Characteristics

Sample Characteristics (N: 28): Recruited from University of Hertfordshire	
Gender	
Male	50%
Female	50%
Age	
<25 (but over 18 years)	7%
26-35	43%
36-45	29%
46-55	11%
Occupation	
Student	39%
Academic/faculty staff	43%
Researcher	18%
Educational/career background	
Technology related	50%
Non technology related (e.g. law)	50%

The Final Phase involved the subjects completing several questionnaires.

The questionnaires relevant to this study were:

a) *Cogniron Introductory Questionnaire:* This enquired about participants' personal details (age, gender, occupation), level of familiarity with robots, prior experience with robots (at work, as toys, in movies/books, in TV shows, in museums or in schools), and level of technical knowledge of robots were rated according to a 5-point Likert scale.

b) *The Subject Personality Questionnaire* – This was based on selected traits from Eysenck's three personality factors [13]: 1) *neuroticism* vs. *emotional stability* (anxiety, tension, shyness, emotional vulnerability), *extraversion* vs. *introversion* (sociability, general activity level, assertiveness, excitement-seeking, dominance), and *psychoticism* (aggressiveness, impulsiveness, creativity). Some of Eysenck's traits were considered unsuitable for self-assessment (e.g. antisocial).

Moreover, we considered only traits that could be rated for both human and robot personality. Autonomy was an additional trait added by our research team. Subjects were required to rate themselves for each of the 13 different personality traits using a 5-point Likert scale (e.g. for autonomy, continuous scale from 1 = prefer being told what to do to 5 = prefer to decide myself what to do).

c) *The Robot Personality Questionnaire*: This questionnaire followed a similar format to the subject personality questionnaire using the 5-point Likert scale, and included the following personality traits from Eysenck’s model: anxiety, tension, shyness, emotional vulnerability, sociability, general activity level, assertiveness, excitement-seeking, dominance, aggressiveness, impulsiveness and creativity. A number of personality characteristics were added to the study and included: autonomy, intentionality, predictability of behaviour, controllability, and considerateness. Robot autonomy was measured from ‘seemed to do what it was told/programmed to do’ to ‘seemed to make its own decisions’.

Robot Behaviour Styles

Subjects were exposed to two different robot behaviour styles and their reactions were recorded for the different situations. A counterbalanced design was used to avoid habituation effects, where the same task scenarios were used to test both robot behaviour conditions [Socially Ignorant (A), Socially Interactive (B)], which were defined a priori by the research team. The robot behaviour styles (A) and (B) constituted different behaviours for the Negotiated Space Task, and the Assistance Task (Tables II & III)

TABLE II: Robot behaviour styles for negotiated space task

Negotiated Space Task	Socially Ignorant (A)	Socially Interactive (B)
Path	Straight line	Circuitous route with respect to subject’s pose
Speed	Fast	Slow when close to subject, most hesitant
Camera	Static & facing forward	Moving & tracking (showing interest in subject’s task)
Encounter	“excuse me”, continue as soon as possible	“After you”, continues after subject has moved away

The robot behaviour styles were always referred to as behaviours A and B to ensure the experimenter did not give away procedural clues. The Socially Ignorant (A) behaviour style was expressed when the robot made little or no change to its behaviour when the participant was present. This corresponds to a robot treating a human not ‘special’ in any way but simply as an obstacle. In contrast, the Socially Interactive (B) behaviour style was classified if the robot took human presence into account by modifying ‘robot optimum behaviour’ (e.g. for a robot to go from x to point y, the ‘optimum’ behaviour in an uncluttered environment is a straight line). A socially interactive robot was thus designed to be ‘considerate’ towards the subject. A mixture of autonomous programmes (e.g. wandering) and Wizard of Oz (WoZ) remote control was used.

TABLE III: Robot behaviour styles for Assistance Task

Assistance Task	Socially Ignorant (A)	Socially Interactive (B)
1st Encounter		
Action	Does not wait for subject	Waits for subject to look/ask for pen
Speech	“I notice you need a pen. I’ll go fetch one”	“I notice you need a pen. I’ll go fetch one”
2nd Encounter		

Action	Bringing basket with pen to side of table, close to subject, putting it down	Waiting in front of table facing subject, waiting for subject to notice, then putting basket down
Speech	“Here you are, please take the pen”	“Here you are, please take the pen”

III. RESULTS

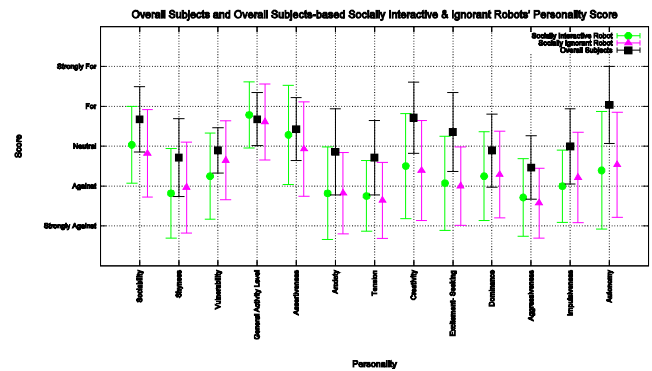
Differences in personality traits between subjects and robots

Paired sample t-tests were computed to determine whether there were any significant differences between subjects’ and robot A and B personality characteristics. Subjects rated themselves as being significantly more sociable, shy, more anxious, tenser, more creative, higher in excitement-seeking, more dominant, more aggressive, and more autonomous compared to both robots A and B. Subjects also rated themselves as being more assertive compared to robot A, more vulnerable compared to robot B, and more impulsive compared to robot B. All t-test results were significant at $p \leq .01^2$. Fig. 3 illustrates means and standard deviations of personality characteristics, for subjects and both robots A & B.

These results seem to suggest that overall, the adult subjects in the current trials did not view their own personality as being similar to the personality characteristics of the robot for both robot behaviour styles (i.e. Socially Ignorant, Socially Interactive). It was interesting that in most cases they felt that they had more ‘extreme’ or stronger personality characteristics compared to the robot for both positive and negative traits. This could be an indicator that they did not view the robot as having particularly prominent or strong personality characteristics. Subjects also did not distinguish between their ratings of the personality between robot A and B, although they displayed different behaviours during the trials towards the two different behaviour styles.

Overall relationship between subject and robot personality

Pearson correlation coefficients were calculated to examine the overall relationship between subject and robot personality traits overall. Only one significant positive linear relationship was revealed between subject dominance and Socially Ignorant robot (A) dominance [$r(26) = 0.44$; $p = 0.02$]. This indicates that the more dominant subjects rated themselves as being on the subject personality questionnaire, the more dominant they perceived robot A as being (or vice versa). No significant correlations were found between subject personality and the Socially Interactive robot (B).



² Due to space limitations, individual t-test values are not shown

Fig. 3 Mean and Standard deviations for personality characteristics for subjects and robots A and B

The impact of gender in relation to subject and robot personality traits

Significant correlations were found between subject personality traits and ratings of robot personality for males and females. For males, the more anxious they rated themselves, the more anxious they rated robot A, and the more psychotic males rated themselves, the more psychotic they rated robot A. For females, a different pattern of findings emerged for links between their own personality and robot personality. The more assertive females rated themselves, the more assertive they rated robot A, and the more dominant they rated themselves, the more dominant robot A was rated. No significant associations were revealed between subject personality and robot B.³

The impact of subject age in relation to subject and robot personality traits

For analysis purposes, two groups were formed, younger subjects were classified as <35 years, and older participants >35 years. Pearson correlation coefficients revealed only one significant negative linear relationship between older subjects and robot A for aggressiveness [r (14) = -0.62, p = 0.02]. This finding indicates that the less aggressive older subjects rated their personality, the more aggressive they rated the personality of robot A.

For younger subjects, a different pattern of findings emerged. Five significant positive linear relationships were found between younger subjects' personality traits and robot A for assertiveness [r (14) = 0.70, p = 0.005], anxiety [r (14) = 0.64, p = 0.02], aggressiveness [r (14) = 0.55, p = 0.04], and impulsiveness [r (14) = 0.57, p = 0.03]. These findings show that the more assertive, anxious, aggressive and impulsive the younger subjects rated themselves, the more assertive, anxious, aggressive and impulsive they perceived the personality of Robot A. No significant relationships emerged for robot B for both older and younger subjects.

Technology related experience in relation to subject and robot personality traits

Subjects' level of technological experience was considered in relation to subject and robot personality traits. Firstly, no significant correlations were found between non-technology background subjects, and their personality ratings and robot personality ratings. In contrast, a number of significant correlations were found between subjects who had a technology related background, their own personality ratings and robot personality ratings. The more anxious and aggressive subjects with a technological background rated themselves, the more aggressive and anxious they rated robot A. A negative linear relationship was found for general activity level, where the less active they perceived themselves, the more active they perceived robot A. Positive linear relationships were revealed for robot B, i.e. the more anxious and excitement-seeking subjects rated themselves, the more excitement-seeking they perceived robot B. A negative linear relationship was found for shyness, where the less shy subjects

with a technology related background rated themselves, the shyer they rated robot B.

Overall, these findings seem to indicate differences between those subjects who had a technology-related background compared to those who did not. Subjects with a technology-related background appeared to relate to both robots A and B on a number of personality traits, suggesting perhaps that they believe that the robot has personality characteristics. However, the fact that no associations were found for the non-technology related subjects seems indicative that they did not view the robot as having any human-like personality characteristics.

The degree of personality attribution between subjects and the robot

To determine the degree to which subjects projected their own personality traits to the robot, the *difference* between the personality traits of each subject and how that same trait was evaluated by the subject to the robot (Socially Ignorant robot and Socially Interactive robot) was calculated. To allow for meaningful comparisons, differences were standardized to the score of the subject. In this way, we were able to obtain the differences for all traits (*j*) and each of the *N* subjects (*i*). We then computed the mean discrepancy, standard error and the 95% confidence intervals (*CI*) across each subject for the 13 personality traits discrepancies and for both robot behaviour styles. The mean discrepancy and the corresponding confidence intervals for the socially ignorant robot were plotted in fig. 4 against those of the socially interactive robot. The diagonal line indicates the positions where the degree of discrepancy between self-evaluation and attribution would be the same for both robot behaviour styles (i.e. $\Delta\bar{P}_j^{sub,robot(ign)} = \Delta\bar{P}_j^{sub,robot(int)}$). Points that fall above the diagonal are characteristics that score relatively high for $\Delta\bar{P}_j^{sub,robot(ign)}$ and low for $\Delta\bar{P}_j^{sub,robot(int)}$.

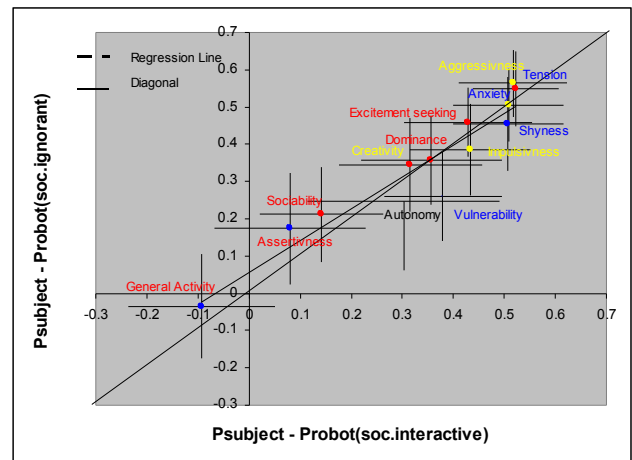


Fig 4. Discrepancies between subject values of personality traits and their attribution to robot behaviour styles A and B. Points are the average discrepancies calculated over measurements of 28 subjects & 95% CI.

Fig. 4 illustrates some interesting features including: 1) The personality traits are relatively close to the diagonal. This indicates that the mean discrepancies measured for the Socially Ignorant robot are linearly related to those of the Socially Interactive robot. 2) The selected traits contributing to

³ Due to space limitations, not all statistical results can be shown

the ‘neuroticism-emotional stability’ and ‘psychoticism’ factor form a cluster and show larger discrepancies (for both robot behaviour styles) than those associated with the ‘extravert-introvert’ dimension. This suggests that subjects evaluated the robot as being more similar to themselves with respect to ‘extra-introvert’ traits than with ‘neuroticism-emotional stability’ and ‘psychoticism’ attributes, 3) Standard deviations are large; hence confidence intervals overlap for many trait discrepancies. This means that only a few of the trait discrepancies differ statistically among each other. Note, only 12 of the 27 traits used by Eysenck were measured in the present study. Further detailed statistical analysis (reported in [14]) suggested the 12 variables should be analysed in their own right, rather than focussing on the original Eysenck factors. Any reference to the ‘Eysenck factors’ in our results therefore strictly relates to the particular traits that we selected (possible alternative factors are proposed in [14]).

IV. DISCUSSION

A summary of the main results of this study revealed that:

- For individual personality traits, subjects perceived themselves as having stronger personality characteristics compared to robots A and B.
- Overall, subjects did not view their own personality as similar to robot behaviours A or B.
- Factors such as subject gender, age and technological experience were important in how subjects viewed their personality as being similar to the robot personality.
- The attribution of personality analysis revealed that subjects evaluated the robot as being more similar to themselves with respect to the traits contributing to the ‘extra-introvert’ factor compared to the ‘neuroticism-emotional stability’ and ‘psychoticism’ factors.

In response to research question one, a number of significant differences were found between subject personality ratings and the personality traits assigned to the robot behaviour styles. This implies that overall, subjects did not view their own personality as being similar to either the Socially Ignorant or Socially Interactive robot behaviour. In most cases, subjects felt that their personality was stronger for both positive and negative personality traits. This could mean that subjects did not view the robots as having a strong or identifiable personality. However, it should be noted that subjects participated in the trials only once within a simulated living-room situation, which does not fully resemble real-life scenarios. It would be interesting in future studies to consider the assignment of personality traits to the robots in more naturalistic surroundings, and longitudinally, to determine whether over time subjects build up a relationship with the robot and start to view it as having a more obvious personality.

Research question two considered the relationship between subject and robot personality. Although overall relationships between subjects’ personality traits and the traits they assigned to the robot did not emerge, factors such as gender, age and technology-based experience were important. In the case of subject gender, relationships were different for males and females, although positive associations were found

in all cases. Males and females appeared to interpret the robot behaviour and personality in different ways. This is an important future design consideration as it suggests that the desired personality and behaviour wanting to be conveyed by the robot may have very different meanings for males and females, and may lead to quite different human-robot interaction styles, and overall satisfaction with the experience.

A different pattern of findings also emerged according to subjects’ age. For older subjects, only one negative relationship was found for aggressiveness ratings of the Socially Ignorant robot. However, for younger subjects, far more significant positive associations were found between subject personality and the personality of robot B. The design implications of this finding suggest that the interaction experiences and interpretations between older and young subjects are very different. This of course could be related to previous exposure to robots, but could imply that older subjects were more anxious and wary of the robot interaction trials compared to younger subjects. If subjects are uncomfortable with interacting with a robot, this could result in them being unfavourable and less engaged towards robots, which might have negative marketing implications for the future of interactive robots, for example robot companions in the home. It is important that future studies examine the impact of age on robot interaction styles more closely and determine whether increased exposure to robots would help to reduce potential anxiety older subjects might have towards robots.

An important finding was that no significant associations were found for those subjects with limited /no technological background compared to a number of associations identified for those subjects from a technology related background. These findings indicate that subjects with no technological knowledge did not view either of the robot behaviour styles as having a personality. This could have important design implications as robot personality traits are likely to assist in human-robot interaction, as it could help the user e.g. to make sense of the interaction, leading to more engaging and believable interactions. Future studies on subjects with non-technological backgrounds could explore the aspects of the robot they find the most and least accepting, and satisfying, and the reasons behind not thinking the robot had personality characteristics. This has implications if future robot companions capable of human-robot interaction are to be accepted by the wider community other than those people with a high interest in robots and technology. However, in the case of robots used e.g. as assistants or toys, it is very likely that people with a technological background will be a strong user group for any new robot product on the market.

The emerging pattern of findings for differences between age and technology related background could link to the argument posed in the introduction that some people may imbue their own personality onto the robot to help them understand, interpret and more fully engage in the interaction with the robot, whereas others may be fearful of losing their own identity and assuming that a robot can have similar personality characteristics and human qualities as them. Young people and those with a technological background seem to be more prepared to assign their own personality traits

onto the robot compared to older subjects and those with a non-technology related background who wish to keep their own personality separate from that of a robot. These findings are related to those reported by Scopelliti et al. [15] that elderly subjects were more frightened at the prospect of having a robot in the home, and showed an element of distrust towards a robot in the home.

A further research question addressed was whether humans projected their own personality onto the robots and whether this depended on the way the robot behaved. Results showed that the degree of attribution of personal characteristics to the robot did not strongly depend on the robot behaviour style (i.e. Socially Interactive or Socially Ignorant) which indicates perhaps that subjects were unable to clearly distinguish the behaviours the robot was exhibiting and related personality characteristics. For example, in the current HRI trials, it could be case that the researchers felt it was polite (i.e. socially interactive) for the robot to wait until the subject looked at the robot in the Assistance Task for the pen, but some subjects commented that they found this irritating, and would therefore have maybe defined the behaviour as being socially ignorant. Future studies should consider in more detail the ability of subjects to distinguish between different types of robot behaviour and related personality constructs as it could help to shape human-robot interaction into a more enjoyable and satisfying experience. This could be achieved e.g. by carrying out detailed pilot studies where subjects use 'Think Aloud' strategies during robot interactions.

To conclude, results from our robot trials indicate that human subjects do not tend to assign their personality traits to match the robots'. This remained the case for different robot behaviour styles. However, subject gender, age and technological background were all important factors related to the extent to which subjects ascribed their own personality traits to the robot. It seems that younger subjects with technology related backgrounds were happy to ascribe their own personality traits to the robot, perhaps in an attempt to understand the interaction more fully. In contrast, older subjects with little technological background did not view their own personality as being similar to that of the robot, perhaps in an attempt to keep their own identity separate to that of the robot. This is a relatively unexplored area of human-robot interaction studies and future research needs to consider the role of robot personality in more detail to fully understand the contribution of personality in emulating engaging human-robot interactions.

ACKNOWLEDGMENTS

We would like to thank Mick Walters, Chrystopher Nehaniv, Iain Werry, David Lee, and all the participants for their involvement in this study.

REFERENCES

- [1] C. L. Breazeal, *Designing sociable robots*. Massachusetts: The MIT Press, 2002.
- [2] T. Minato, M. Shimada, H. Ishiguro, & S. Itakura, "Development of an android robot for studying human-robot interaction," *Lecture Notes in Computer Science*, vol. 3029, pp. 424-434, 2004.
- [3] S. Woods, K. Dautenhahn, & J. Schulz, "The design space of robots: Investigating children's views," *Proc. Ro-Man, Japan*, 2004, pp. 47-52.
- [4] D. C. Dryer, "Getting personal with computers: How to design personalities for agents," *Applied Artificial Intelligence*, vol. 13, pp. 273-295, 1999.
- [5] D. A. Norman, "How might people interact with agents," *Communications of the ACM*, vol. 37, pp. 68-71, 1994.
- [6] S. Woods, K. Dautenhahn, & J. Schulz, "Child and adults' perspective on robot appearance," *Proc. AISB 05, UK*, 2005, pp. 126-132.
- [7] J. Goetz & S. Kiesler, "Cooperation with a robotic assistant," presented at CHI'02, New York, USA, 2002.
- [8] J. Cassell & J. Bickmore, "Negotiated collusion: Modeling social language and its relationship effects in intelligent agents," *User Modeling and User-Adapted Interaction*, vol. 13, pp. 89-132, 2003.
- [9] P. T. Costa & R. R. McCrae, "Personality stability & its implications for clinical psychology," *Clinical Psych. Review*, vol. 6, pp. 407-423, 1986.
- [10] H. J. Eysenck & S. B. G. Eysenck, *Manual of the Eysenck personality questionnaire*. London: Hodder and Stoughton, 1975.
- [11] H. J. Eysenck, "Dimensions of personality: 16, 5 or 3? Criteria for a taxonomic paradigm," *Personality & Ind Differences*, vol. 12, pp. 773-790, 1991.
- [12] P. T. Costa & R. R. McCrae, "Four ways five factors are basic," *Personality & Ind. Differences*, vol. 13, pp. 653-665, 1992.
- [13] H. J. Eysenck & M. W. Eysenck, *Personality & ind. differences: A natural science approach*. New York, USA: Plenum, 1985.
- [14] M. L. Walters, K. Dautenhahn, R. te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. L. Nehaniv, D. Lee, I. Werry, "The Influence of Subjects' Personality Traits on Personal Spatial Zones in a Human-Robot Interaction Experiment", *Proc. IEEE Ro-man*, 2005, pp. 347-352.
- [15] M. Scopelliti, et al., "If I had a robot at home. Peoples' representation of domestic robots," in *Designing a more inclusive world*, S. Keates, et al. Eds. Cambridge, UK: Springer, 2004, pp. 257-266.