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

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.
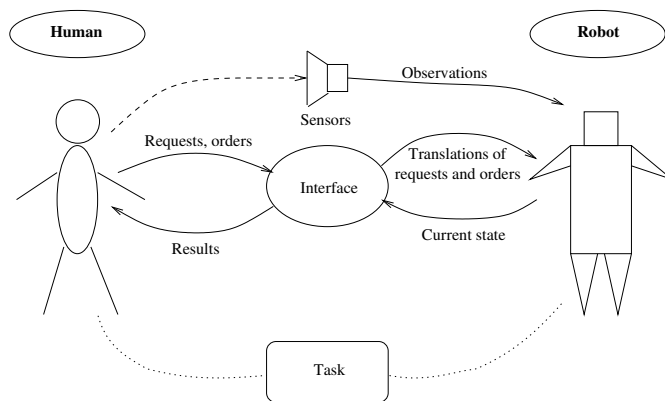

Fig. 1.   General description of HRI

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 to 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).

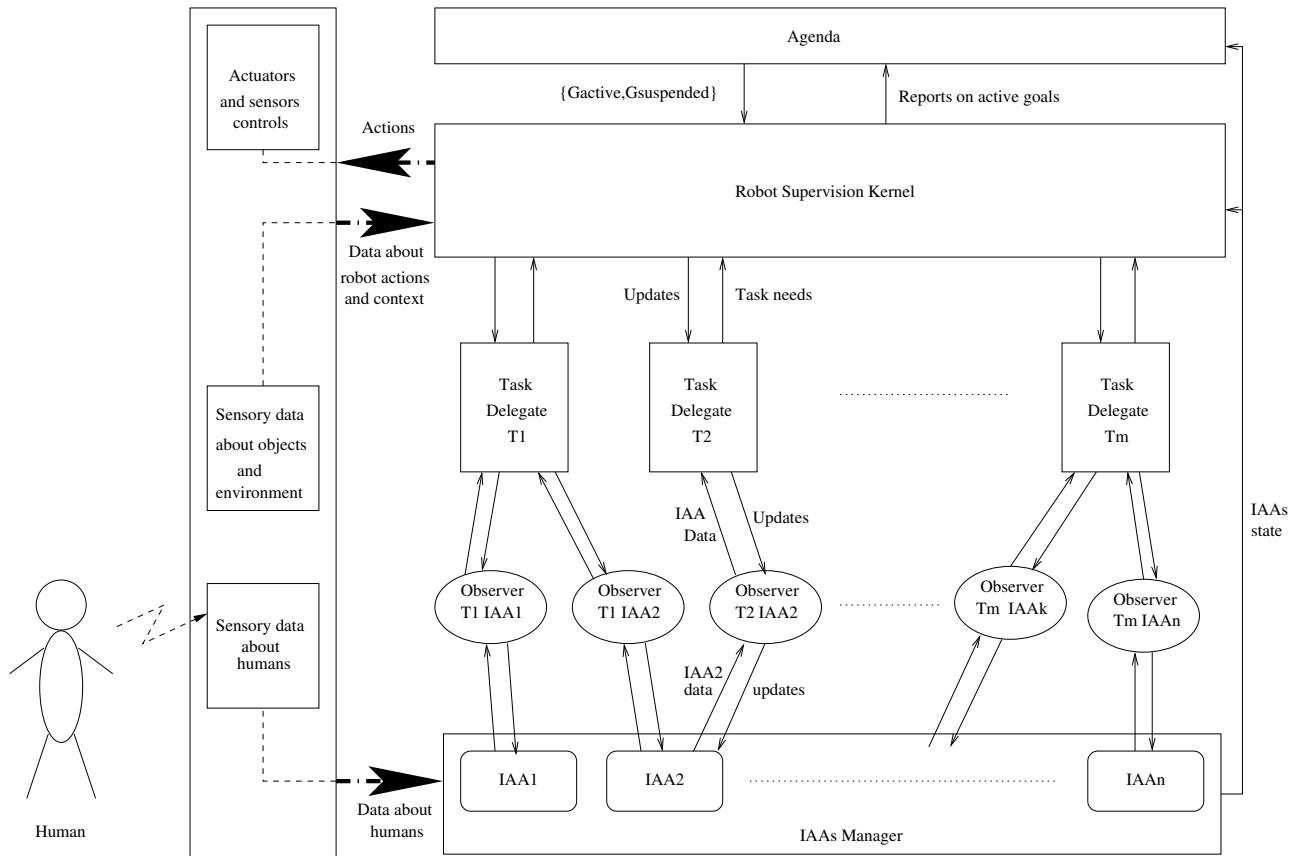The next section provides a more detailed explanation of

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 *recharging*, *search_for_an_IAA*, *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
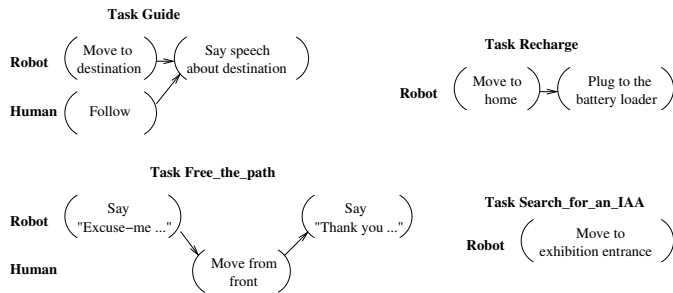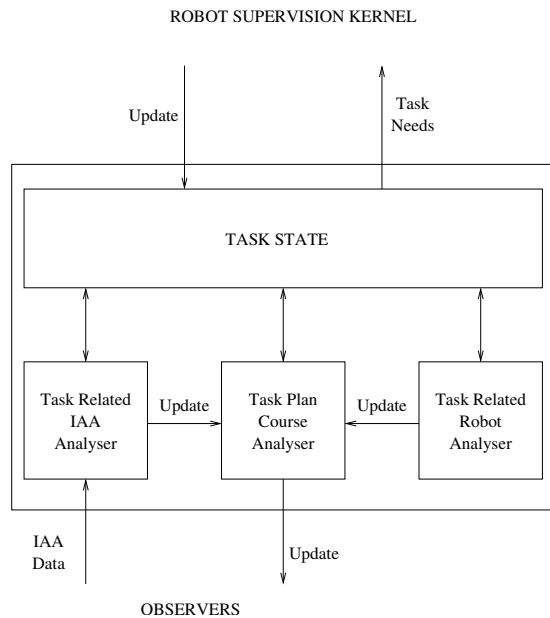
Fig. 4.   Scripts for Rackham goals

ROBOT SUPERVISION KERNEL

Fig. 5.   Structure of a Task Delegate

Fig. 6.   Task Delegate states

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).
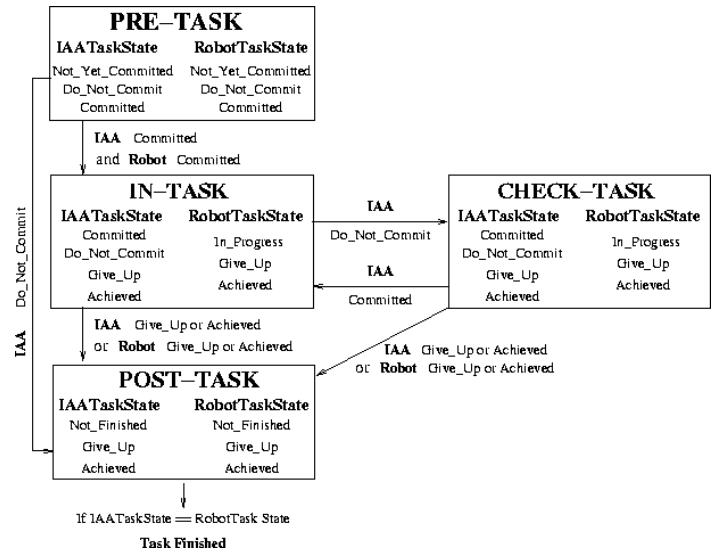
Regarding the progress of the task, the Task Delegate will update the observation rules of OBs and data stored in the Plan Course Analyzer thus permitting synchronization between team members. Given the task course and the context, the Task Delegate can decide to stay in IN-TASK or to switch to CHECK-TASK 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. Bï¿½ar, 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. Kleinehagenbrock, 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 RAS/RSJ 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.