THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Intelligence Artificielle

Présentée et soutenue par :

Mme SANDRA DEVIN

le vendredi 3 novembre 2017

Titre :

Decisional issues during Human-Robot Joint Action

Ecole doctorale :

Mathématiques, Informatique, Télécommunications de Toulouse (MITT)

Unité de recherche :

Laboratoire d'Analyse et d'Architecture des Systèmes (L.A.A.S.)

Directeur(s) de Thèse :

M. RACHID ALAMI

M. MALIK GHALLAB

Rapporteurs :

Mme VANESSA EVERS, UNIV. TWENTE ENSCHEDE PAYS-BAS

M. PETER FORD DOMINEY, INSERM RHONE ALPES - AUVERGNE

Membre(s) du jury :

M. RAJA CHATILA, UNIVERSITE PIERRE ET MARIE CURIE, Président

M. MALIK GHALLAB, LAAS TOULOUSE, Membre

Mme AURELIE CLODIC, LAAS TOULOUSE, Membre

Mme JULIE SHAH, MASSACHUSSETS INSTITUTE OF TECHNOLOGY, Membre

M. RACHID ALAMI, LAAS TOULOUSE, Membre

Remerciements

Je tiens tout d'abord à remercier les deux personnes qui ont participé activement à l'encadrement de ma thèse, Rachid ALAMI et Aurélie CLODIC. Ils ont toujours su me donner des conseils avisés et ont été un soutien sans faille durant ces trois années.

Je remercie Peter FORD DOMINEY et Vanessa EVERS, tout deux rapporteurs de mon manuscrit de thèse, pour m'avoir fait l'honneur d'accepter d'évaluer mon travail et de me faire bénéficier de leur expertise. Je remercie également Raja CHATILA, Julie SHAH et Malik GHALLAB pour avoir accepté de siéger en tant qu'examinateur au sein mon jury.

Je souhaite remercier les autres doctorants, ingénieurs et stagiaires de l'équipe HRI avec qui j'ai pu travailler pour leurs idées, leur support et la bonne ambiance au sein du groupe. Merci Jules, Grégoire, Michelangelo, Harmish, Renaud, Raphaël, Camille, Yoan, Amandine, Mamoun, Erwan, Théo et Jean-François. Je remercie également Aurélie, Matthieu et Jérôme pour avoir répondu à mes nombreux appels à l'aide durant mes manipulations sur le robot.

Je remercie Amélie, Ellon, Nicolas, Laurent, Arthur, Arthur, Christophe, Rafael, Alexandre, Alejandro pour leur bonne humeur et la bonne ambiance apportée au laboratoire et en dehors ainsi que pour les parties de tarot.

Je remercie le CETAM, mon club de judo, pour leur soutien sans faille sur le tatami et en dehors sans lequel je ne serais sûrement pas arrivé au bout de cette thèse. Merci Greg, Etienne, Claire, Cyril, Olivier, Eric, Hugues, Baptise, Mimine, Jérôme, Edwin, Micka, Rachel, Yasushi et tous les autres.

Je remercie mes parents qui ont su faire de moi ce que je suis aujourd'hui.

Je tiens à remercier plus particulièrement Etienne, pour son soutien dans la vie de tous les jours durant ces trois années et ses nombreuses relectures.

Enfin, je remercie Arnaud, pour son soutien sur la fin de la thèse, son amour et bien sûr, pour l'eau du jury.

Abstract

In the future, robots will become our companions and co-workers. They will gradually appear in our environment, to help elderly or disabled people or to perform repetitive or unsafe tasks. However, we are still far from a real autonomous robot, which would be able to act in a natural, efficient and secure manner with humans. To endow robots with the capacity to act naturally with human, it is important to study, first, how humans act together. Consequently, this manuscript starts with a state of the art on joint action in psychology and philosophy before presenting the implementation of the principles gained from this study to human-robot joint action. We will then describe the supervision module for human-robot interaction developed during the thesis. Part of the work presented in this manuscript concerns the management of what we call a shared plan. Here, a shared plan is a a partially ordered set of actions to be performed by humans and/or the robot for the purpose of achieving a given goal. First, we present how the robot estimates the beliefs of its humans partners concerning the shared plan (called mental states) and how it takes these mental states into account during shared plan execution. It allows it to be able to communicate in a clever way about the potential divergent beliefs between the robot and the humans knowledge. Second, we present the abstraction of the shared plans and the postponing of some decisions. Indeed, in previous works, the robot took all decisions at planning time (who should perform which action, which object to use...) which could be perceived as unnatural by the human during execution as it imposes a solution preferentially to any other. This work allows us to endow the robot with the capacity to identify which decisions can be postponed to execution time and to take the right decision according to the human behavior in order to get a fluent and natural robot behavior. The complete system of shared plans management has been evaluated in simulation and with real robots in the context of a user study. Thereafter, we present our work concerning the non-verbal communication needed for human-robot joint action. This work is here focused on how to manage the robot head, which allows to transmit information concerning what the robot's activity and what it understands of the human actions, as well as coordination signals. Finally, we present how to mix planning and learning in order to allow the robot to be more efficient during its decision process. The idea, inspired from neuroscience studies, is to limit the use of planning (which is adapted to the human-aware context but costly) by letting the learning module made the choices when the robot is in a "known" situation. The first obtained results demonstrate the potential interest of the proposed solution.

Resumé

Les robots sont les futurs compagnons et équipiers de demain. Que ce soit pour aider les personnes âgées ou handicapées dans leurs vies de tous les jours ou pour réaliser des tâches répétitives ou dangereuses, les robots apparaîtront petit à petit dans notre environnement. Cependant, nous sommes encore loin d'un vrai robot autonome, qui agirait de manière naturelle, efficace et sécurisée avec l'homme. Afin de doter le robot de la capacité d'agir naturellement avec l'homme, il est important d'étudier dans un premier temps comment les hommes agissent entre eux. Cette thèse commence donc par un état de l'art sur l'action conjointe en psychologie et philosophie avant d'aborder la mise en application des principes tirés de cette étude à l'action conjointe homme-robot. Nous décrirons ensuite le module de supervision pour l'interaction homme-robot développé durant la thèse.

Une partie des travaux présentés dans cette thèse porte sur la gestion de ce que l'on appelle un plan partagé. Ici un plan partagé est une séquence d'actions partiellement ordonnées à effectuer par l'homme et/ou le robot afin d'atteindre un but donné. Dans un premier temps, nous présenterons comment le robot estime l'état des connaissances des hommes avec qui il collabore concernant le plan partagé (appelées états mentaux) et les prend en compte pendant l'exécution du plan. Cela permet au robot de communiquer de manière pertinente sur les potentielles divergences entre ses croyances et celles des hommes. Puis, dans un second temps, nous présenterons l'abstraction de ces plan partagés et le report de certaines décisions. En effet, dans les précédents travaux, le robot prenait en avance toutes les décisions concernant le plan partagé (qui va effectuer quelle action, quels objets utiliser...) ce qui pouvait être contraignant et perçu comme non naturel par l'homme lors de l'exécution car cela pouvait lui imposer une solution par rapport à une autre. Ces travaux vise à permettre au robot d'identifier quelles décisions peuvent être reportées à l'exécution et de gérer leur résolutions suivant le comportement de l'homme afin d'obtenir un comportement du robot plus fluide et naturel. Le système complet de gestions des plan partagés à été évalué en simulation et en situation réelle lors d'une étude utilisateur.

Par la suite, nous présenterons nos travaux portant sur la communication non-verbale nécessaire lors de de l'action conjointe homme-robot. Ces travaux sont ici focalisés sur l'usage de la tête du robot, cette dernière permettant de transmettre des informations concernant ce que fait le robot et ce qu'il comprend de ce que fait l'homme, ainsi que des signaux de coordination. Finalement, il sera présenté comment coupler planification et apprentissage afin de permettre au robot d'être plus efficace lors de sa prise de décision. L'idée, inspirée par des études de neurosciences, est de limiter l'utilisation de la planification (adaptée au contexte de l'interaction homme-robot mais coûteuse) en laissant la main au module d'apprentissage lorsque le robot se trouve en situation "connue". Les premiers résultats obtenus démontrent sur le principe l'efficacité de la solution proposée.

Contents

Introduction

Contents

Context

In the 1940s, researchers invented the first machines that we can call computers. Then, they quickly came to think that this new tool which can easily manipulate numbers can also manipulate symbols and they started to work on new "thinking machines". In 1956, at the Dartmouth conference, the domain of "Artificial Intelligence" is recognized as a fully academic field. Associated to the automaton technology, the first "robots" quickly arrived in our environment.

Some of these robots are meant to work alone (e.g. rovers for space exploration) while others need to work in the vicinity and/or with humans. One possible example is robot "co-workers". These robots need to collaborate in a safe, efficient and fluent way with humans to accomplish more or less repetitive tasks. The last decades also witnessed the advent of what is called "sociable robots" [Dautenhahn 2007]. These robots can be used, for example, to help elderly or disabled people in their daily life or to guide people in public spaces.

The aim of this thesis is to make a step toward robots which act jointly with humans in a natural, efficient and fluent way. We focus more especially on the decisional issues that can appear during human-robot Joint Action. The subject of Joint Action between humans has been studied extensively in social sciences, however, many things remain to be discovered. Based on these results, the aim here is to build robots which are able to understand the humans' beliefs and choices, and to adapt to them in order to be more pleasant and efficient companions.

Human-Robot Joint Action challenges

Constructing robots which are able to smoothly execute Joint Action with humans brings a number of challenges.

A first prerequisite for the robot is to be engaged in the Joint Action. If the goal is not imposed by its human partner, the robot needs to pro-actively propose its help whenever it is needed. Then, the robot needs to monitor its partner engagement in the task and exhibit its engagement in the same task.

Once the robot has a goal to achieve, it needs to be able to find a plan to achieve this goal. This plan should be feasible in the current context of course, but it should also take the human into account, his abilities and preferences. Once a plan found, the robot should be able to share it or negotiate it with its partner. Only then, the execution of the task can begin.

During the Joint Action execution, one first challenge for the robot is to be able to understand how the humans perceive their environment and what is their knowledge concerning the task. In other words, it needs to be able to constantly estimate the mental states of its partners. These mental states should be taken into account at every steps of the execution in order to ensure a good understanding between partners.

Finally, the robot needs to be able to coordinate with the human. This coordination is needed at all level of the execution. At a lower level, the robot should exhibit an understandable and predictable behavior when performing actions. It also needs to be able to execute actions such as handover which require precise motor coordination with its partner. At a higher level, the robot needs to coordinate the Shared Plan execution. It not only needs to execute its actions at the right time but it should also give the appropriate information at the right time to its partners using either verbal or non-verbal communication.

Contributions and manuscript organization

At the beginning of the document, I add two main starting points. One of these points was, as in almost all thesis I guess, the current states of the art both in robotics and in social sciences concerning human-human Joint Action. The second starting point was the current architecture for human-robot interaction developed in our research group and more especially the supervision part which I was asked to develop it. The aim with all of this being to bring innovating changes to the current system (again as in most thesis I guess). The process which I followed during this thesis and which I will explain now can be graphically summarized in Fig. 1.

As a first step, I studied the literature concerning Joint Action between humans in order to better understand what are the needed components of a successful Joint Action. I also studied the current state of the art in robotics, and more especially in human-robot interaction, in order to have an overview of what robots were already capable of. Based on all of this, I identified the needed components of a successful Human-Robot Joint Action and I studied, always based on bibliography, how to

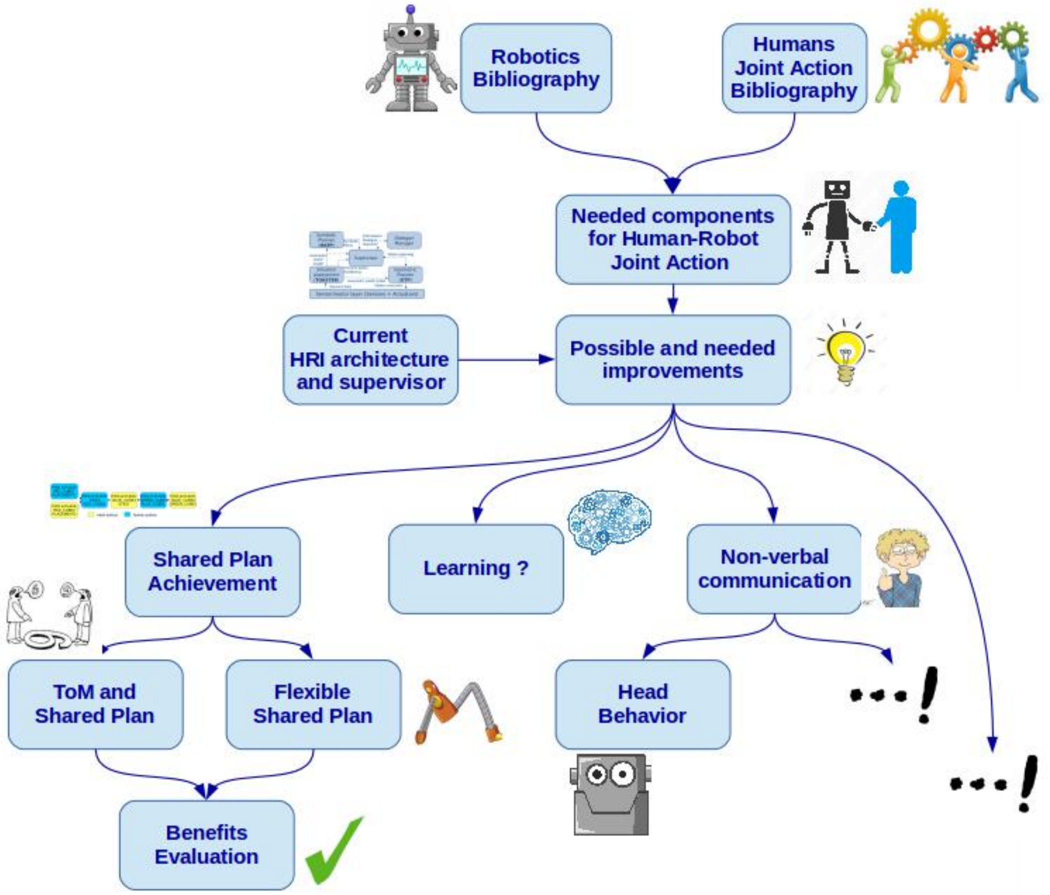


Figure 1: Organization of the contributions presented in the manuscript.

articulate them into a coherent architecture. This part constitutes the Chapter 1 of the manuscript.

Then, I took a look at the current state of the architecture for human-robot interaction developed in our research group. I especially focused on the supervision part of the architecture which I was in charge to develop. Based on the conclusion of my bibliographic study, I was able to identify several possible improvements to bring to the supervisor. The final version of the supervisor is presented in Chapter 2 in order to help the understanding of the following of the manuscript and constitute the major technical contribution of the thesis.

One first subject where I saw possibilities of improvement was the way the robot elaborates and executes Shared Plans. This topic is discussed in the second part of the manuscript and is decomposed in three chapters:

— First, I noticed that there was a gap between the perspective abilities of the robot and the Shared Plan execution. Indeed, several previous works endowed the robot with the ability to estimate how its human partners perceive the world and how to use this knowledge on several domains such as dialogue. However, there was no work to link this ability to the Shared Plan execution. In Chapter 3, I will present how I endowed the robot with the ability to estimate the humans mental states, not only about the environment, but also concerning the state of the task and more particularly of the Shared Plan. Then, I will present how the robot is able to use these mental states to better communicate about divergent beliefs during Shared Plan execution.

— Then, based on discussions with psychologists on the subject, we noticed that the way the robot was dealing with Shared Plans was not "natural" for humans and not flexible enough. Indeed, in the previous system, the robot was taking all decisions during Shared Plan elaboration. It was choosing for each action who should perform it and with which specific objects. Imagine a table with several identical objects on it that the robot needs to clean in collaboration with a human. The robot would have decided at the beginning for each object who should take it and in which exact order. The human would have simply removed objects and adapts to the robot decisions. Chapter 4 aims to reduce this gap. We first identified the needed decisions during Shared Plan elaboration and execution and we endowed the robot with the ability to decide which decisions should be taken at planning time and which one could better be postponed at execution time. Then, we allowed the robot to take these decision by smoothly adapting to the human choices.

— Finally, we wanted to evaluate both improvements brought to the Shared Plan achievement by the robot. We did it in Chapter 5 both quantitatively in simulation and qualitatively with a user study with the real robot. These studies allowed to show the pertinence of the proposed improvements and to compare two different modes developed in the context of this work (in one mode the robot negotiates some needed decisions during Shared Plan execution while in the other the robot adapts to the human choices). Moreover, for the purpose of the user study, a questionnaire has been developed to evaluate the users feelings concerning the collaboration with the robot. This questionnaire has been validated (in term of inter coherence) thanks to the study data and is generic enough to be considered as a future tool for human-robot collaboration evaluation.

Then, I focused on another interesting work subject concerning the non-verbal behavior of the robot. Indeed, during Joint Action between humans, Joint Action participants exchange a lot of information through non-verbal communication. It allows to increase fluency in the task execution and to align the knowledge of all participants. Consequently, for the robot to become a better Joint Action partner, it should be able to provide such information with its non-verbal behavior. In Chapter 6, we studied more especially the head behavior of the robot (there is plenty other ways to give information with non-verbal behavior, but it may more need a career than a thesis to study all of them). Based on the bibliography in social sciences and on previous works in robotics, we identified needed components of a robot head behavior adapted to the Joint Action. We studied more deeply some of them with an on-line video based study. To conclude this chapter, we present how these components can be implemented into a robot head behavior architecture.

Finally, in the context of the RoboErgoSum ANR project [1], I have been brought to work in collaboration with ISIR at Paris where researchers focus on learning for robot high level decision. In Chapter 7, with another PhD student of ISIR Erwan Renaudo, we studied how to combine planning and learning in the context of human-robot Joint Action. The idea is to take advantage from both sides in order to come up with decision level which is able to quickly learn how to smoothly adapt to the human choices during Joint Action execution.

Work environment

This thesis has been realized at LAAS-CNRS in the RIS team (Robotics and InteractionS). It was included in the general objective to build a robotics architecture for an autonomous robots which interacts with humans.

Robot: in all this thesis, for practical reasons, the developed algorithms have been implemented in a PR2 robot from Willow Garage [2]. However, these algorithms are generic enough to be implemented in other robots. The PR2 robot is a semi-humanoid robot which is able to navigate and manipulate objects (see Fig.2).

Figure 2: The PR2 robot.

1. http://roboergosum.isir.upmc.fr/
2. http://wiki.ros.org/Robots/PR2

Humans and objects detection: When interacting with humans during manipulation tasks, the robot needs to be able to localize and identify humans and objects. To avoid as much as possible perception issues which are not the focus of this thesis, the perception of humans and objects is simplified here. The humans are identified and perceived thanks to a motion capture system. They wear a helmet to get the position and orientation of their heads and a glove to get the position and orientation of their right hands (see Fig. 3). Concerning the objects, they are identified and localized with tags thanks to the robot cameras in its head.

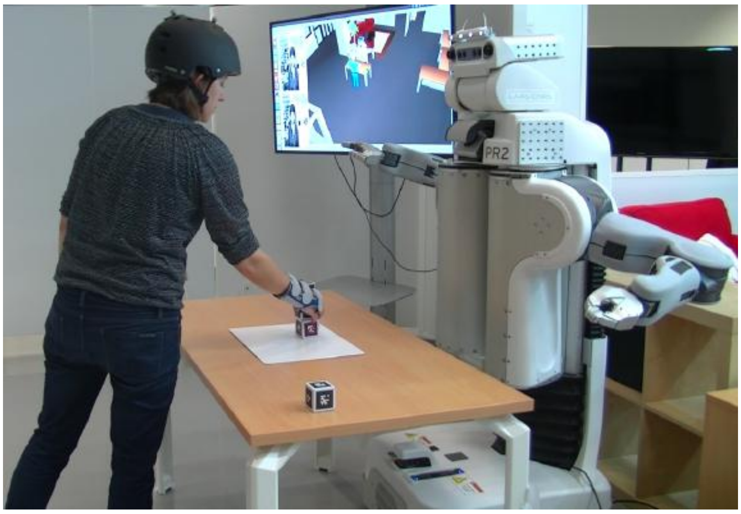


Figure 3: The PR2 robot interacting with a human to build a stack of cubes. The human is detected thanks to a motion capture system (helmet and glove) and the objects with tags.

Contributions

Publications

The work presented in this thesis has led to several publications. They are listed here below (from the most recent to the oldest):

— **Devin, S.**, Clodic, A., Alami, R. (2017). *About Decisions During Human-Robot Shared Plan Achievement: Who Should Act and How?* The Ninth International Conference on Social Robotics (ICSR).

— **Devin, S.**, Alami, R. (2016). *An implemented theory of mind to improve human-robot shared plans execution.* In Human-Robot Interaction (HRI), 2016 11th ACM/IEEE International Conference on (pp. 319-326). **Best Paper.**

— **Devin, S**., Milliez, G., Fiore, M., Clodic, A., Alami, R. (2016). *Some essential skills and their combination in an architecture for a cognitive and interactive robot.* Workshop In Human-Robot Interaction (HRI), 2016 11th ACM/IEEE International Conference on (pp. 319-326).

— Khamassi, M., Girard, B., Clodic, A., **Devin, S.**, Renaudo, E., Pacherie, E., Alami, R., Chatila, R. (2016). *Integration of Action, Joint Action and Learning in Robot Cognitive Architectures.* Intellectica Journal, 2016(65), 169-203.

— Renaudo, E., **Devin, S.**, Girard, B., Chatila, R., Alami, R., Khamassi, M., Clodic, A. (2015). *Learning to interact with humans using goal-directed and habitual behaviors.* In RoMan 2015, Workshop on Learning for Human-Robot Collaboration. IEEE.

Other

In addition to publications, several other scientific contributions have been made to the field:

— **Devin, S.**, Clodic, A., Alami, R. (2017). *Shared Plans for Human-Robot Joint Action.* Talk at the 7th Joint Action Meeting (JAM7).

— **Devin, S.**, Clodic, A., Alami, R. (2015). *A Theory of Mind for Human-Robot Joint Action.* Talk at the 6th Joint Action Meeting (JAM6).

— *Towards a Framework for Joint Action.* Co-organization of a series of workshops on Joint Action between philosophers, psychologists and roboticists. Joint Action Meeting 7 (JAM 2017), RO-MAN (2016), ICSR (2015).

— **Devin, S.**, Alami, R. (2016). *An implemented Theory of Mind to improve Human-Robot Shared Plans execution.* Talk at the "Journée de travail sur la robotique interactive et cognitive" of the GDR robotics.

Part I

From Human-Human Joint Action to a supervisor for Human-Robot Interaction

Chapter 1

From Human-Human Joint Action to Human-Robot Joint Action

Contents

1.1 Joint Action Theory

A first step to endow robots with the ability to perform Joint Actions with humans is to understand how humans act together. As a working definition of Joint Action, we will use the one from [Sebanz 2006]:

> *Joint action can be regarded as any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment.*

A given number of prerequisites are needed for these individuals to achieve the so-called Joint Action. First of all, they need to agree on the change they want to bring in the environment, the conditions under which they will stay engaged in its realization and the way to do it. A number of works have studied this topic, relative to *commitment*, which I will develop in Sec. 1.1.1. Then, as mentioned in the definition, the individuals need to coordinate their actions in space and time.

This will be studied in Sec. 1.1.3. Finally, in order to coordinate, each individual needs to be aware of the other, he needs to be able to perceive him and predict his actions. It will be developed in Sec. 1.1.2.

1.1.1 Commitment

The first prerequisite to achieve a Joint Action is to have a *goal* to pursue and the *intention* to achieve it. Let's define what is called a *goal* and an *intention* for a single person before going to a *joint goal* and a *joint intention.*

In [Tomasello 2005], Tomasello et al. define what they call a *goal* and an *intention* and illustrate these definitions with an example and an associated figure (fig. 1.1) where a person wants to open a box.

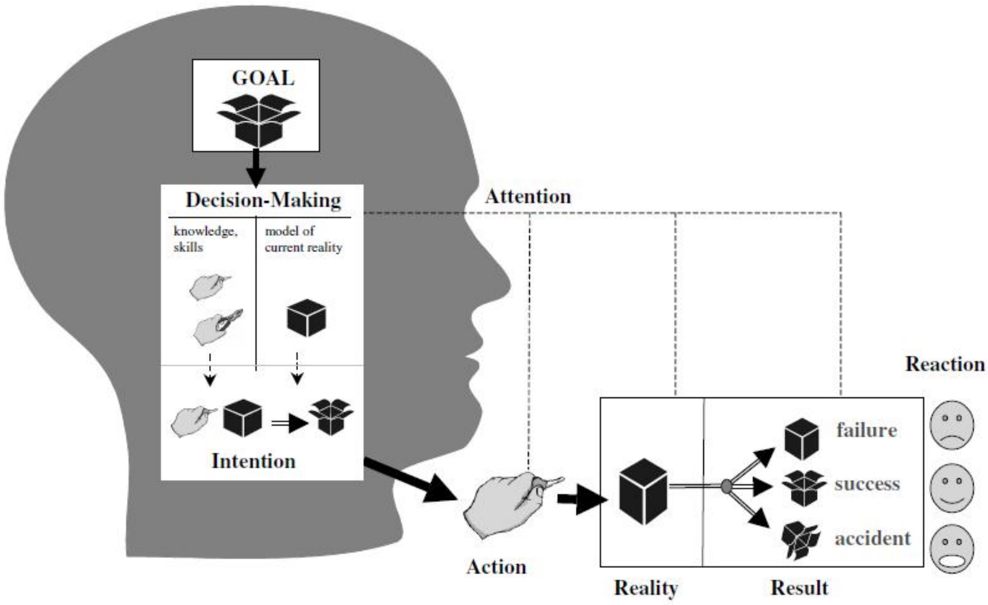


Figure 1.1: Illustrative example of an intentional action by Tomasello et al. Here the human has the *goal* for the box to be opened. He chooses a means to perform it and so forms an *intention.*

A *goal* is defined here as the representation of the desired state by the agent (in the example, the goal is an open box) and, based on Bratman's work [Bratman 1989], an *intention* is defined as an action plan the agent commits to in pursuit of a goal (in the example, the intention is to use a key to open the box). The *intention* includes both a *goal* and the means to achieve it.

Cohen and Levesque propose in [Cohen 1991] a formal definition of what they call a *persistent goal* relative to a condition q (i.e. the goal is considered valid only if q is true):

Definition: An agent has a *persistent goal* relative to q to achieve p iff:

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 neither be true, or that q is false.

However, their definition of an *intention* differs a little from the previous one. They define an *intention* as a commitment to act in a certain mental state:

> **Definition:** An agent *intends* relative to some conditions to do an action just 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.

The *intention* still includes the *goal* but here it concerns more the fact that the agent commits to achieving the goal than the way to achieve it.

Let's now apply these principles to a Joint Action. One of the best known definition of *joint intention* is the one of Bratman [Bratman 1993]:

> We intend to J if and only if:
>
> 1. (a) I intend that we J and (b) you intend that we J.
> 2. I intend that we J in accordance with and because of $1a$, $1b$, and meshing subplans of $1a$ and $1b$; you intend that we J in accordance with and because of $1a$, $1b$, and meshing subplans of $1a$ and $1b$.
> 3. 1 and 2 are common knowledge between us.

This definition is taken back and illustrated by Tomasello et al. in [Tomasello 2005] where they reuse the example of the box to open (fig D.1).

The *shared goal* is defined as the representation of a desired state plus the fact that it will be done in collaboration with other person(s) (in the example, they will open the box together) and a *joint intention* is defined as a collaborative plan the agents commit to in order to achieve the *shared goal* and which takes into account both agents individual plans (here an agent will hold the box with the clamp while the other open it with the cutter).

In a same way, Cohen and Levesque extend their definition of *persistent goal* and *intention* to a collaborative activity. They first define a *weak achievement goal* as:

> **Definition:** An agent has a *weak achievement goal* relative to q and with respect to a team to bring about p if either of these conditions holds:
>
> — The agent has a normal achievement goal to bring about p, that is, the agent does not yet believe that p is true and has p eventually being true as goal.
>
> — The agent believes that p is true, will never be true, or is irrelevant (that is, q is false), *but* has as a goal that the status of p be mutually believed by all the team members.

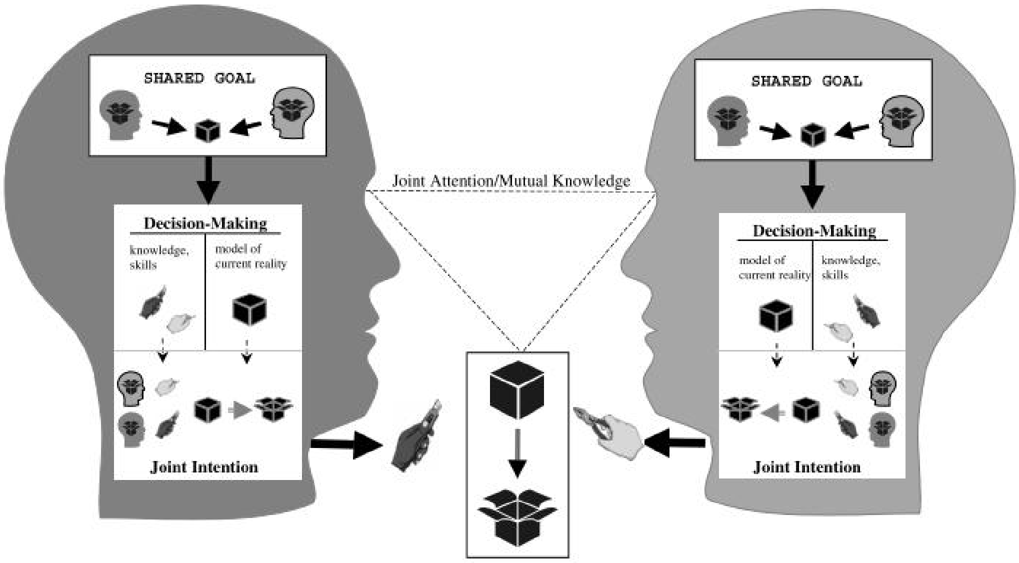


Figure 1.2: Illustrative example of a collaborative activity by Tomasello et al. Here the humans have for *shared goal* to open the box together. They choose a means to perform it which takes into account the other capabilities and so form a *joint intention*.

They then use this definition to define a *joint persistent goal*:

> **Definition:** A team of agents have a *joint persistent goal* relative to q to achieve p just in case
> — they mutually believe that p is currently false;
> — they mutually know they all want p to eventually be true;
> — it is true (and mutual knowledge) that until they come to mutually believe that p *is true*, that p will never be true, or that q is false, they will continue to mutually believe that they each have p as a weak achievement goal relative to q and with respect to the team.

They finally define a *joint intention* as:

> **Definition:** A team of agents *jointly intends*, relative to some escape condition, to do an action iff the members have a joint persistent goal relative of that condition of their having done the action and, moreover, having done it mutually believing throughout that they were doing it.

As previously, the definitions of Cohen and Levesque do no take into account the way to achieve the *shared goal*, however, they introduce the interesting idea that agents are also engaged to inform each other about the state of the *shared goal*.

Concerning the way to achieve a *shared goal*, mentioned into the definition of the *joint intention* of Tomasello et al., Grosz and Sidner initially introduce and formalize the notion of *Shared Plan* in [Grosz 1988], which is extended in [Grosz 1999]. The key properties of their model are:

1. it uses individual intentions to establish commitment of collaborators to their joint activity
2. it establishes an agent's commitments to its collaborating partners' abilities to carry out their individual actions that contribute to the joint activity
3. it accounts for helpful behavior in the context of collaborative activity
4. it covers contracting actions and distinguishes contracting from collaboration
5. the need for agents to communicate is derivative, not stipulated, and follows from the general commitment to the group activity
6. the meshing of subplans is ensured it is also derivative from more general constraints.

With their definition, each agent does not necessarily know the whole *Shared Plan* but only his own individual plan and the meshing subparts of the plan. The group has a *Shared Plan*, but no individual member necessarily has the whole *Shared Plan*.

In conclusion, the concepts concerning the commitment of agents to a collaborative activity that we will use in this thesis can be summarized as:

— A *goal* will be represented as a desired state.

— A*shared goal* will be considered as a *goal* to be achieved in collaboration with other partner(s). An agent is considered engaged in a *shared goal* if he believes the goal is currently false, he wants the goal to be true and he will not abandon the goal unless he knows that the goal is achieved, not feasible or not relevant any more and he knows that his partners are aware of it.

— A *joint intention* will include a *shared goal* and the way to realize it, represented as a *Shared Plan* which will take into account the capacities of each agent and the potential conflicts between their actions. This *Shared Plan* will not be necessarily completely known by all members of the group but all individuals will know their part of the plan and the meshing subparts.

If we apply this to the box example of Tomasello it gives us:

— "The box will be open" can be a *goal* for an agent.

— "The box will be open because we collaborate" can be a *shared goal* for several agents. Once the agents agree to achieve this goal, they will not give up until the box is open (and the other agent knows it), the box can not be opened (and the other agent knows it) or there is no more need to open the box (and the other agent knows it).

— A joint intention for two agents relative to the *shared goal* to open the box will be, for example, that the first agent go get the opener, he gives it to the second agent and then the second agent opens the box with the opener. The sequence of actions <go get the opener, give it, open the box> is the Shared Plan. The second agent does not need details concerning the part "go get the opener" while the first agent does not to know need details concerning the part "open the box".

1.1.2 Perception and prediction

One important thing for an agent when performing a Joint Action is to be able to perceive and predict the actions of his partner and their effects. Based on the works in [Sebanz 2006], [Pacherie 2011] and [Obhi 2011] we identified several necessary abilities for this predictions:

Joint attention: The capacity for an agent to share focus with his partner allows to share a representation of objects and events. It brings a better understanding of the other agent's knowledge and where his attention is focused and so, it helps the prediction of his possible next actions. Moreover, there should be a mutual manifestation of this joint attention, meaning that we should show that we share the other attention.

Action observation: Several studies have shown that when someone observes another person executing an action, a corresponding representation of the action is formed for the observer [Rizzolatti 2004]. This is done by what has been called the *mirror-neuron* system. This behavior allows the observer to predict the outcomes of the actor's action.

Co-representation: An agent needs to have a representation of his partner, including his goal, his capacities and the social rules he is following. This representation also includes the knowledge of the partner on the Shared Plan, especially on the actions attributed to him. Having this representation will help to predict his future actions. For example, as a pedestrian knows that the car drivers follow the traffic regulations, he will be able to predict that they will stop if he sees a red traffic light.

Agency: Sometimes, when there is a close link between an action performed by oneself and an action performed by someone else, it can be hard to distinguish who caused a particular effect. The capacity to attribute the effects to the right actor is called the sense of *Agency*. This sense of *Agency* is important in Joint Action in order to correctly predict the effects of each action.

Based on the same works mentioned above and on [Sebanz 2009], we can list several kinds of predictions to support Joint Action which can be done thanks to the abilities described previously :

— **What:** A first one is to predict what an agent will do. Two kinds of predictions, described in [Pacherie 2011], can be distinguished here:
  — *action-to-goal:* this is supported by the *mirror-neuron* system introduced before. Here the word goal designates the goal of an action, its purpose.

The idea is that by observing an action, it is possible to predict its goal. For example, if we observe someone extending his arm toward an object we can predict that he will pick the object.
 — *goal-to-action:* here the word goal designates the goal of a task, as defined in the previous subsection. Knowing this goal, it can be easy to predict which action an agent will perform in a given context.
— **When:** another prediction which is necessary is the timing of an action. Knowing when an action will occur and how long it will take allows for a better coordination in time.
— **Where:** a Joint Action usually takes place in a shared space. It is therefore necessary to predict the future position of the partner and his actions in order to coordinate in space or reason about affordances of the other agent.

1.1.3 Coordination

The predictions discussed previously allow agents to coordinate during Joint Action. Two kinds of coordination are defined in [Knoblich 2011] that both support Joint Action.

Emergent coordination: It is a coordinated behavior which occurs unintentionally, independently of any joint plan or common knowledge and due to perception-action couplings. Four types of sources of emergent coordination can be distinguished:
— *Entrainment:* Entrainment is a process that leads to temporal coordination of two actors' behavior, in particular, synchronization, even in the absence of a direct mechanical coupling. It is the case, for example, for two people seating in rocking chairs involuntary synchronizing their rocking frequencies [Richardson 2007].
— *Affordances:* An object affordance represents the opportunities that an object provides to an agent for a certain action repertoire [Gibson 1977]. For example, the different ways to grab a mug. Two kinds of affordances can lead to an emergent coordination: *common affordances* and *joint affordances*. When several agents have the same action repertoire and perceive the same object they have a *common affordance*. This *common affordance* can lead the agents to execute the same action. When an object has affordances for two or more peoples collectively, the agents have to synchronize for an action to occur. This is what is called *joint affordances*. For example, a long two-handled saw affords cutting for two people acting together but not for either of them acting individually.
— *Perception-action matching:* As discussed before, observing an action activates corresponding representation in the observer's mind. This process can lead to involuntary mimicry of the observed action. Consequently, if two persons observe the same action, they can have the same reaction to mimic the action.

— *Action simulation:* The internal mechanisms activated during action observation not only allow to mimic the action but also to predict the effects of this action. If two people observe the same action and so predict the same effects, they can consequently have the same reaction. For example, two persons seeing the same object falling will have the same reaction to try to catch it.

Planned coordination: While emergent coordination is unintentional, planned coordination requires for agents to plan their own actions in relation to Joint Action and others' actions.

One way for an agent to intentionally coordinate during Joint Action is to change his behavior compared to when he is acting alone. These changes of behavior are called *coordination smoothers* in [Vesper 2010] and can be of several types:

— Making our behavior more predictable by doing for example wider or less variable movements
— Structuring our own task in order to reduce the need of coordination. For example sharing the space or working turn by turn.
— Producing coordination signals like looking someone who should act or counting down.
— Changing the way we use an object by using an affordance more appropriate to a shared use.

Another way to coordinate is through communication. Indeed, Clark argues that two or more persons cannot perform a Joint Action without communicating [Clark 1996]. Here the word communication includes both verbal and non-verbal communication. Clark also defines what he calls the *common ground*: when two agents communicate, they necessarily have common knowledge and conventions. Moreover, when communicating, it is important to not only send a message but also to assure that the message has been understood as the sender intends it to be. This process to make the sender and the receiver mutually believe that the message has been understood well enough for current purposes is called *grounding*.

In conclusion, in order to smoothly perform a Joint Action, an agent needs to:

— Develop sufficient perception and prediction abilities in order to coordinate in space and time. This needs to be done from basic motor commands to high level decisions.
— Produce coordination signals understandable by his partners in order for them to predict his behavior.
— Ensure that the signals he sends are well received by his partners.

1.2 How to endow a robot with Joint Action abilities

In this section we will discuss how the theory on human-human Joint Action can be applied to human-robot Joint Action. Following what has been discussed on commitment, we will first see in Sec. 1.2.1 how the robot can engage in Joint Action and understand the intention of its human partners. Then, we will see in Sec. 1.2.2 how the robot perceives the humans and can predict their actions by taking into account their perspectives and mental states. We will also see how the robot can coordinate during Joint Action in Sec. 1.2.3.

The topics which are linked to the work presented in this thesis will be more developed in the corresponding chapters.

1.2.1 Engagement and Intention

Goal management: As for humans, robots need to be able to engage in Joint Action. A first prerequisite is to choose a goal to perform. This goal can be imposed by a direct order of the user, however, the robot also needs to be able to pro actively propose its help whenever a human needs it. To do so, the robot needs to be able to infer high-level goals by observing and reasoning on its human partners' activities. This process is called plan recognition or, when a bigger focus is put on human-robot interaction aspects, intention recognition. Many works have been done concerning plan recognition using approaches such as classical planning [Ramırez 2009], probabilistic planning [Bui 2003] or logic-based techniques [Singla 2011]. Concerning intention recognition, works such as [Breazeal 2009] and [Baker 2014] take into account theory of mind aspects to deduce what the human is doing.

When direct orders have been received and humans intentions recognized, the robot needs to choose which goal to perform, also taking into account its own resources. This problem has not been addressed as a whole in the literature, however, some similar works can be seen as partial answers. For example, some deliberation systems allow to solve problems with multiple goals taking into account resources such as time [Georgeff 1987, Ghallab 1994, Lemai 2004] or energy level [Rabideau 1999]. In AI, the goal reasoning domains deals with some similar problems [Molineaux 2010, Roberts 2016]. The role of goal reasoning is to survey the current goals of a robot, check that they remain feasible and relevant and establish new goals if needed.

Once the robot is engaged in a Joint Action, it needs to be able to monitor other agents engagement. Indeed, it needs to understand if, for a reason, a human aborts the current goal and reacts accordingly. This can be done using gaze cues and gestures [Rich 2010], postures [Sanghvi 2011] but also context and humans mental states [Salam 2015].

Shared Plan management: Finally, once a goal is chosen, a Shared Plan needs to be established for the robot and its human partners to achieve the goal. Several works have been done in task planning to take into account the human [Cirillo 2010, Lallement 2014]. They allow the robot to reduce resource conflicts [Chakraborti 2016], take divergent beliefs [Warnier 2012, Talamadupula 2014] or socio-emotional state [Charisi 2017] into account , or promote stigmergic collaboration for agents in co-habitation [Chakraborti 2015]. Once the plan computed, the robot needs to be able to share/negotiate it with its partners. Several studies have been reported on how to communicate these plans. Some researchers studied how a system could acquire knowledge on plan decomposition from a user [Mohseni-Kabir 2015] and how dialog can be used to teach new collaborative plans to the robot and to modify these plans [Petit 2013]. In [Pointeau 2014], a cognitive system combining episodic and semantic memory allows the robot to learn relations between objects as well as temporal relations forming the basis for memories of Shared Plans. In [Sorce 2015], the system is able to learn a plan from a user and transmit it to another user and in [Allen 2002] a computer agent is able to construct a plan in collaboration with a user. Finally, [Milliez 2016b] presents a system where the robot shares the plan with a level of details which depends on the expertise of the user.

1.2.2 Perspective taking and humans mental states

One of the first difference between a human and a robot is the way they perceive the world. To perceive its environment, the robot uses sensors to recognize and localize entities. These sensors return positions and orientations in the form of coordinates (x, y, z, θ, ϕ, ψ). On the other hand, humans use relations between objects to describe their positions (e.g. the mug is on the kitchen table, oriented toward the window). To understand the human references and to generate understandable utterances, the robot needs therefore to build a semantic representation of the world, based on the geometric data it collects from sensors. This process is called *grounding* and has been developed in several works [Coradeschi 2003, Mavridis 2005, Lemaignan 2012].

However having its own semantic representation of the world is not enough for the robot, it also needs to take into account the point of view of its partners in order to better understand their goals and actions. To do so, the robot does what is called *perspective taking*, it constructs a representation of the world from the humans perspectives [Breazeal 2006, Milliez 2014]. This ability can be used by the robot to choose its actions in order to influence others mental states [Gray 2014], solve ambiguous situations [Ros 2010a] or to better interact during dialogue [Ferreira 2015].

One important application of *perspective taking* is human action recognition. Indeed, knowing what others are aware of is a first step to understand what they are doing. Then, the action recognition can be done based on Partially Observed Markov Decision Processes (POMDP) and Dynamic Bayesian Networks (DBN) [Baker 2014] or inverse reinforcement learning [Nagai 2015].

This subject will be more developed and we will see, in Chapter 3, how we use *perspective taking* to estimate mental states of the humans concerning the Shared Plans in order to improve their execution.

1.2.3 Coordination

One of the most important and difficult challenges during human robot Joint Action is to coordinate. The problem appears at different levels during Joint Action.

At a higher level, the humans and the robot need to coordinate their actions to fluently execute the Shared Plan. The robot needs to execute its actions at the right time, monitor the ones of its partners and correctly communicate when needed. Several systems have been developed to do so, as *Chaski* [Shah 2011], a task-level executor which uses insights from human-human teaming in order to minimize human idle time or *Pike* an online executive that unifies intention recognition and plan adaptation to deal with temporal uncertainties during Shared Plan execution [Karpas 2015]. A part of the work presented in the thesis is the extension of *SHARY*, a supervisor allowing to execute Shared Plans into a complete human-aware architecture [Clodic 2009]. We will notably see in Chapter 4 how we extend it in order to execute flexible Shared Plans where part of the decisions are postponed until the execution.

One of the key aspects at this level of coordination is verbal and non-verbal communication. Concerning verbal communication, there are two ways to consider dialogue. The first one consists on seeing dialogue as a Joint Action. The second one is to see it as a tool for Joint Action. In practice, dialogue can be both, and, as developed in [Clark 1996], there can be Joint Actions in Joint Actions. Several works in robotics developed modules allowing the robot to perform dialogue with humans in support to Joint Action [Roy 2000, Lucignano 2013, Ferreira 2015]. To support dialogue and Joint Action in general, non verbal communication is very important. Its benefit has been shown for human-robot interaction [Breazeal 2005] and ways to perform it have been studied, mainly concerning gaze cues [Boucher 2010, Mutlu 2009b] but also postures [Hart 2014]. However, there are few works which study the use of non-verbal behavior during human-robot Joint Action where both partners are acting. This subject and the associated literature will be more developed in Chapter 6.

At a lower level, the robot needs to coordinate with its partners during action execution. To execute a task, the robot can be led to perform actions in collaboration with one or several humans. The principal action studied in HRI is handover, an action which seems simple as we do it in every day life but which, in fact, raises a number of challenges as, among others, approaching the other person [Walters 2007], finding an acceptable posture to give the object [Cakmak 2011, Dehais 2011, Mainprice 2012] or releasing the object with the good timing [Mason 2005]. But, the robot also needs to coordinate when it is executing an action on its own. Indeed, it needs to share space and resources and its actions need to be understandable enough for its partners. To do so, the robot not only has to execute its actions in an efficient way, but also in a legible, acceptable and predictable way. This process can be compared to the *coordination smoothers* described in [Vesper 2010] and one way to do it is through human-aware motion planning [Sisbot 2012, Kruse 2013].

1.3 A three levels architecture

We saw previously the different prerequisites to Joint Action, both between humans and in HRI. We will see now how the monitoring of Joint Action is organized around three different levels first with the theory of Pacherie concerning humans Joint Actions in Sec. 1.3.1 and then in the LAAS robotics architecture in Sec. 1.3.2.

1.3.1 The three levels of Pacherie

As for the prerequisite of Joint Action, we will first introduce the concepts developed by Pacherie on Action and then, extend them to Joint Action. Pacherie is a philosopher and argues in [Pacherie 2008] that intention in Action is composed of three levels which all have a specific role to play and which are organized as in fig. D.2.

Distal Intention: This is the highest level of intention. In a first time, this level is in charge of forming an intention to act. It means that it is in charge of choosing a goal, a time to execute it and finding a plan to achieve it. Then, once the time comes to execute the plan, this level has to ensure its good execution. To do so, Pacherie takes back the definition of what is called *rational guidance and control* [Buekens 2001]. This control takes two forms: 'tracking control' where we ensure that each successive step in the action plan is successfully implemented before moving to the next step and 'collateral control' where we control for the side effects of accomplishing an action.

Proximal Intention: This level inherits an action plan from the *Distal Intention*. Its responsibility is, first, to anchor the received action plan which is defined in an abstract way in the situation of the action. It needs to integrate conceptual information about the intended action inherited from the *Distal Intention* with perceptual information about the current situation to yield a more definite representation of the action to be performed. Then, this level has to ensure that the imagined actions become current through situational control of their unfolding.

Motor Intention: This is the lowest level of intention. As for the other levels, it first has to make choices and then to monitor their executions. At this level, these choices concern motor commands, which are the physical ways to achieve the action inherited from the *Proximal Intention*.

In [Pacherie 2011], Pacherie extends these three levels to Joint Action. In the same way as before, these three new levels coexist at the same time, each one controlling the Joint Action at a different level.

Shared Distal Intention: Where *Distal Intention* was responsible for intention, *Shared Distal Intention* is responsible for joint intention. When performing a Joint

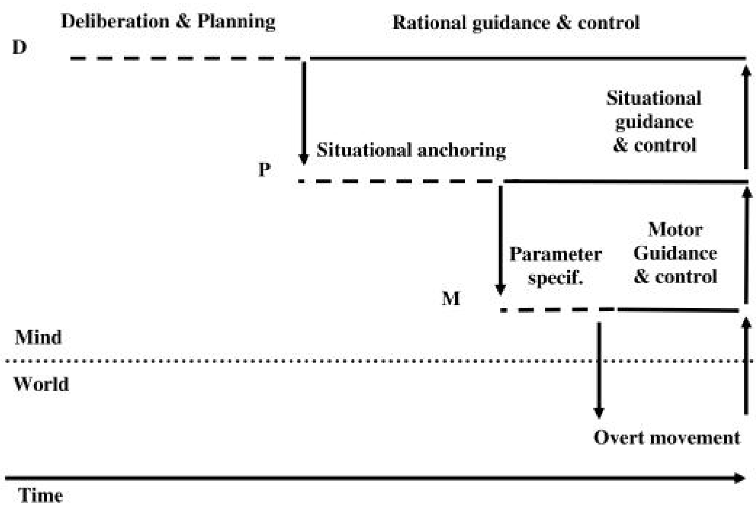


Figure 1.3: The intentional cascade by Pacherie. Distal, Proximal and Motor intentions coexist at the same time, each one controlling the action at a different level. Image from [Pacherie 2008].

Action, this level is the one responsible for the shared goal and the Shared Plan. As said in Sec. 1.1.1, the agent does not have a whole representation of the Shared Plan here and part of his representation will be executed by someone else.

Shared Proximal Intention: This level has the same responsibilities as *Proximal Intention*, however, the anchoring of the action plan needs to take care of the Joint Action partners and to be done in a coordinated way. During the monitoring part, the choices made previously need to be adapted to the others' behavior.

Coupled Motor Intention: As for *Motor Intention*, this level is responsible for the motor commands of the agent. During Joint Action, this level will be the one responsible for precise spatio-temporal coordination for the actions which need it (e.g. holding an object together).

1.3.2 A three levels robotics architecture

Ten years before Pacherie came with her action theory with three levels, the field of autonomous robotics was trying to build architectures and was already intuitively designing three similar levels. A first implemented architecture for autonomous robots is presented in [Alami 1998], organized around these three levels (fig. D.21).

Decision level: This level can be compared to the *Distal Intention* level of Pacherie. It is the one responsible for producing a task plan and supervising it. It sends actions to execute and receives reports from the *execution level*.

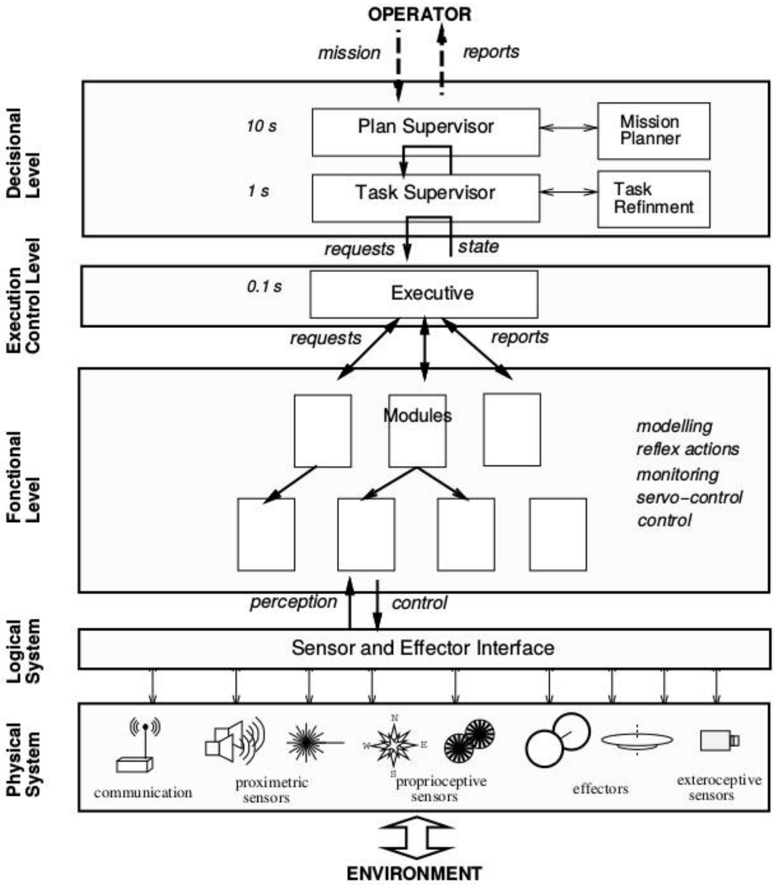


Figure 1.4: One of the first architectures for autonomous robots. The architecture is divided in three main parts: decision, execution and functional levels. Image from [Alami 1998].

Execution level: This level can be compared to the *Proximal Intention* level of Pacherie. It receives from the *decision level* the sequence of actions to be executed and selects, parametrizes and synchronizes dynamically the adequate functions of the *functional level*.

Functional level: This level can be compared to the *Motor Intention* level of Pacherie. It includes all the basic robot action and perception capacities (motion planning, vision, localization, tracking motion control...).

In the past years, this architecture has been developed and adapted to the field of HRI. In recent works, we presented in [Devin 2016b] a theoretical version of the architecture adapted to human-robot Joint Action and still based on the three levels of Pacherie (fig. D.3). The implemented version of this architecture will be presented in Chapter 2, where my contribution in the architecture will also be highlighted.

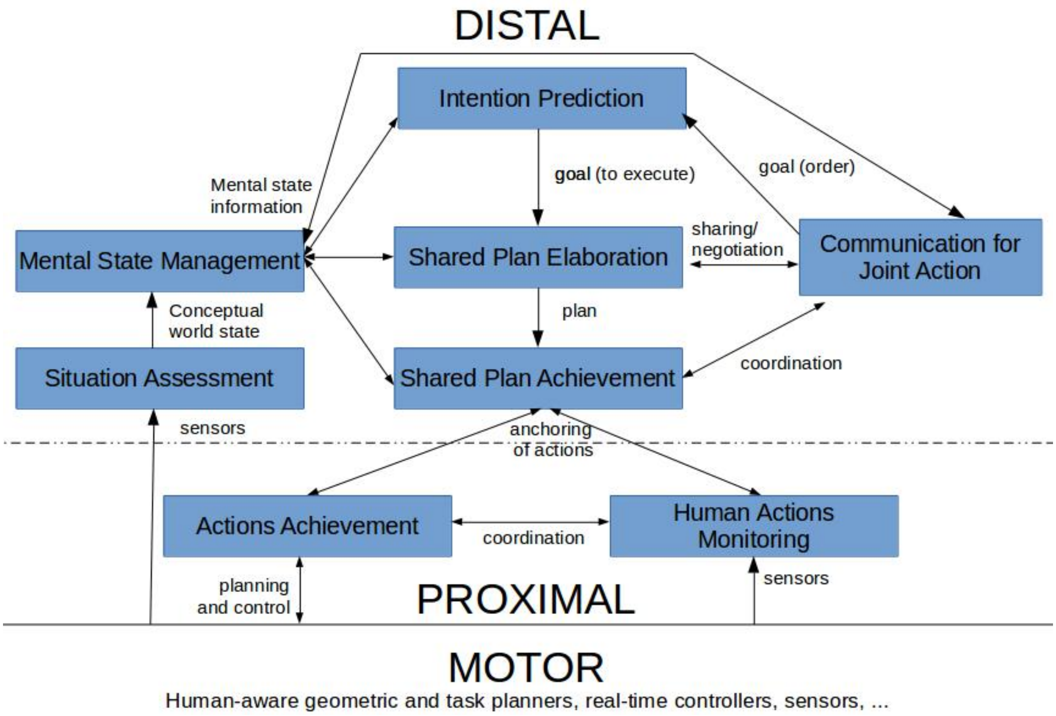


Figure 1.5: Recent architecture for human-robot Joint Action. The architecture is organized in three levels corresponding to the ones defined by Pacherie.

Distal level: As for *Shared Distal Intention*, this level is responsible for goals and Shared Plans management. At this level, the robot is supposed to reason on its environment with high level representations. To do so, the robot is equipped with a **Situation Assessment** module which builds a symbolic representation of the robot environment. To be able to also reason about the humans knowledge, the robot is equipped with a **Mental State Management** module which constantly estimates humans mental states. With this information, the **Intention prediction** module is able to estimate humans intention and if the robot should propose its help or not. This module determines the goal of the robot and allows, during its execution, to monitor other agents engagement. Once the goal chosen, the **Shared Plan Elaboration** module allows the robot to construct and negotiate a Shared Plan to achieve the goal. Then, the **Shared Plan Achievement** module monitors the good execution of this Shared Plan. The last part of this level is the **Communication for Joint Action** module which allows the robot to verbally and non-verbally communicate during Joint Action.

Proximal level: As for *Shared Proximal Intention*, this level is in charge of anchoring the Shared Plan actions in the current situation. This level is composed of two parts: the **Actions Achievement** module which allows to call the adequate motor modules at the right time in order to perform robot actions and the **Human**

Actions Monitoring module which allows to recognize and interpret humans actions with regard to the Shared Plan. These two modules communicate in order to coordinate robot actions to the humans ones.

Motor level: As for *Coupled Motor Intention*, this level is in charge of motor commands of the robot. This level includes all modules allowing to control the robot actuators and interprets data from sensors. These modules, or at least a part of them, also take into account the humans as, for example, the human-aware geometric task and motion planner.

1.3.3 Comparison to other robotics architectures

We saw previously that human-robot interaction is a very complex field with many interesting subjects to study. As a consequence, few architectures allow the robot to execute tasks with humans in a fully autonomous way.

In [Baxter 2013], Baxter et al. present a cognitive architecture built around DAIM (The Distributed Associative Interactive Memory), a memory component which allows the robot to classify the humans behavior. This architecture allows the robot to fluently align its behavior with the human while memorizing data on the interaction. However, several key aspects of human-robot Joint Action as human-aware action execution or theory of mind are missing in this architecture.

Another cognitive architecture is presented in [Trafton 2013]. ACT-R/E, a cognitive architecture based on the ACT-R architecture, is used for human-robot interaction tasks. The architecture aims at simulating how humans think, perceive and act in the world. ACT-R/E has been tested in different scenarios, such as theory of mind and hide and seek, to show its capacity of modeling human behaviors and thought. This architecture has a big focus on the theory of mind and decisional aspects letting less space to the human-aware action execution or understanding, which are also important HRI challenges.

[Beetz 2010] proposes a cognitive architecture called CRAM (Cognitive Robot Abstract Machine) that integrates KnowRob [Tenorth 2013], a knowledge processing framework based on Prolog. CRAM is a very complete architecture dealing with problems such as manipulation, perception, plans or beliefs management. However, this architecture is more designed for a robot acting alone than a robot acting in collaboration with a human. Consequently, the architecture misses some key Joint Action aspects such as communication or humans actions monitoring.

An architecture based on inverse and forward models is presented in [Demiris 2006]. This architecture integrates interesting mechanisms which use human point of view to achieve learning. However, these aspects are limited to action recognition and execution and the architecture does not allow to deal, in its current implementation, with higher level decisional issues.

All of these architectures are really interesting and sharp in their respective predilection domains. However, even if our architecture may lack of some aspects as learning or memory management, its aim is to integrate a major part of the human-robot Joint Action aspects from higher level (intention management and decisional process) to lower level (human-aware execution, coordination and perception). Moreover, the architecture has been conceived in a modular enough way to allow the addition of new modules.

Chapter 2

Supervision for Human-Robot Interaction

Contents

2.1 Role of the supervisor in the global architecture

One of the goals of our research group at LAAS-CNRS is to build a fully autonomous robot which interacts and performs Joint Actions with humans. To do so, an architecture for human-robot interaction has been developed and is constantly improved by the group. This architecture is composed of several modules and a simplified scheme of it can be found in Fig. D.4.

Sensorimotor layer: The lower level of the architecture is composed of modules which allow to communicate and control sensors and actuators. Among others, this layer is composed of modules interpreting sensors data to detect humans and objects and a module allowing to execute given trajectories by calling the adequate actuators.

Situation Assessment: The situation assessment is done by a software called TOASTER [Milliez 2016a]. One of the functionalities of TOASTER is to build and maintain a consistent world state based on data coming from the sensorimotor layer. Geometric computations are done on this world state to compute symbolic

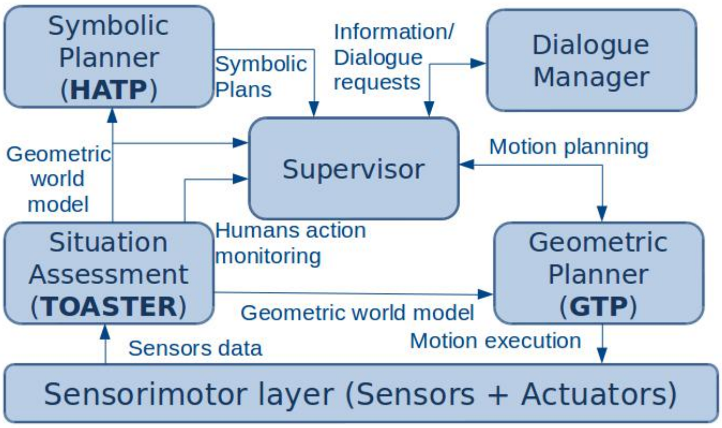


Figure 2.1: The global architecture for human-robot interaction implemented at LAAS-CNRS.

predicates concerning the environment (e.g. *<object, isOn, support>*, *<object, isIn, box>*) and agents abilities and behaviors (e.g. *<object, isVisibleBy, human>*, *<object, isReachableBy, robot>*, *human, isLookingToward, object*). TOASTER is also in charge of perspective taking: the previous predicates are constantly estimated and maintained not only from the robot's point of view but also from the point of view of all humans concerned and perceived in the current context. All the data concerning the world state are stored and accessible through a database.

Geometric Planner: In order to perform actions and movements adapted to the human proximity, our architecture is equipped with a geometric task and motion planner called GTP [Waldhart 2016]. GTP allows to compute trajectories as well as objects placements and grasp. It does that at a level that is human understandable and readable by giving access to high level tasks such as Pick or Place while taking into account the human safety and comfort.

Symbolic Planner: For the robot to be able to synthesize Shared Plans, our architecture is equipped with HATP (Human-Aware Task Planner), a human-aware HTN (Hierarchical Task Network) task planner which allows the robot to compute and refine a plan both for itself and its humans partners, taking into account a number of social rules [Lallement 2014]. HATP has been specially designed to integrate a number of features that are meant to promote the synthesis of plans that are acceptable by humans and easily if not trivially understandable by them. It allows to specify the humans and robot capabilities in terms of actions they can execute. Several aspects such as human preferences and comfort, estimation of human effort to achieve a task in a given context and "social rules" are used in a cost-based approach to build "sufficiently good" human-robot Shared Plans. An example of a plan computed by HATP can be found in Fig. 2.2.

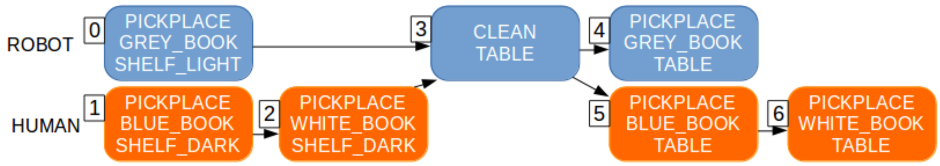


Figure 2.2: An example of a Shared Plan computed by HATP. This plan allows a human and a robot to clean a table by removing all objects on it, cleaning it and then putting back all previous objects.

Dialogue Manager: In order for the robot to communicate with humans, a basic dialogue manager has been integrated to the architecture. This module allows to give humans information concerning the environment (it verbalize predicates), ask basic questions (as asking if a human want to perform an action) and understand basic answers (mainly yes or no answers, the user can answer with buttons as there is no speech recognition in the system).

Supervisor: The last module of the architecture is the supervisor. It is the one in charge of controlling collaborative activities. It chooses the robot goals and monitors the Shared Plan execution. To do so, it estimates humans mental states concerning the Shared Plan and takes them into account to decide when to perform actions or to communicate (verbally and/or non-verbally). It also interprets the information coming from the Situation Assessment module in order to recognize human actions like Pick or Place with regard to the Shared Plan. This module is an extension of [Clodic 2009] and [Fiore 2016] and is the major technical contribution of this thesis. Its internal architecture will be detailed in the next section.

2.2 The supervisor architecture

The supervisor is composed of several modules and is fully implemented in ROS [1]. The complete scheme of its architecture can be found in Fig. D.5, however, since the figure is quite complex, each composing part of the supervisor will be represented and described individually in the next subsections.

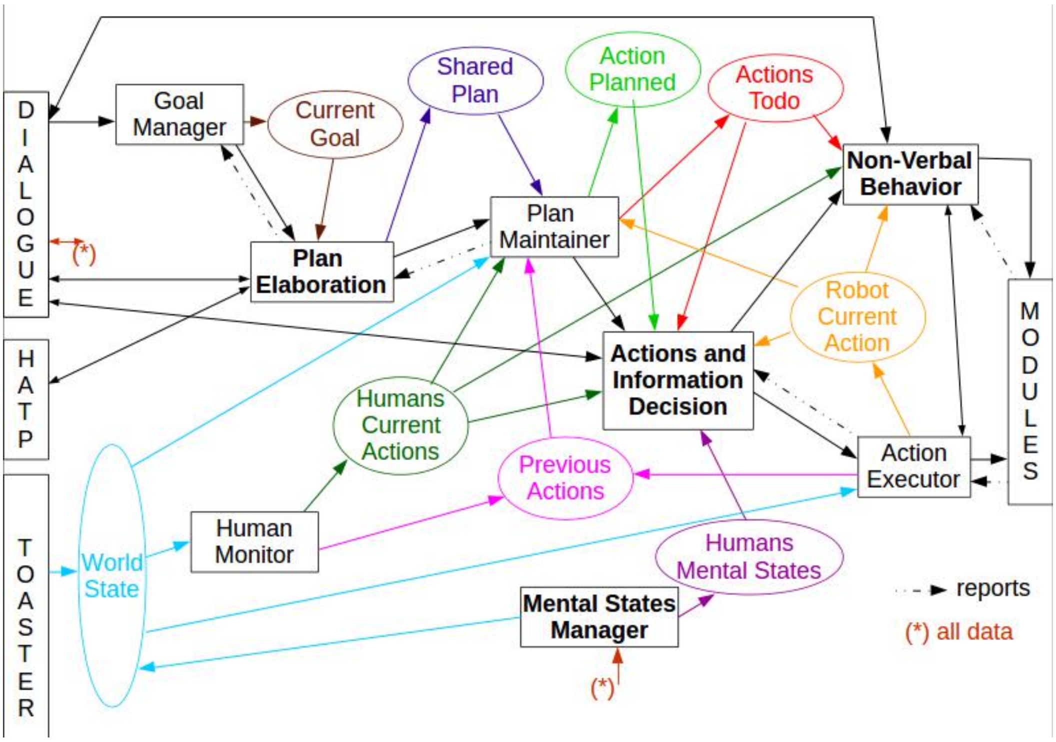


Figure 2.3: Architecture of the supervisor. The modules in bolt are the ones developed in this manuscript.

2.2.1 Goal Manager

The **Goal Manager** allows the robot to select and prioritize goals. It maintains a priority list of goals to perform. This list is updated with *insert*, *abort* or *halt* commands from dialogue or command line.

The chosen goal is published in order for the **Plan Elaboration** module to compute a Shared Plan to satisfy it. The **Goal Manager** sends *stop* and *suspend* orders to the **Plan Elaboration** from which it receives reports concerning the success of the plan or the impossibility to find a plan.

This module is, for now, really basic. An interesting extension would be to integrate data coming from an intention recognition module concerning humans

1. http://www.ros.org/

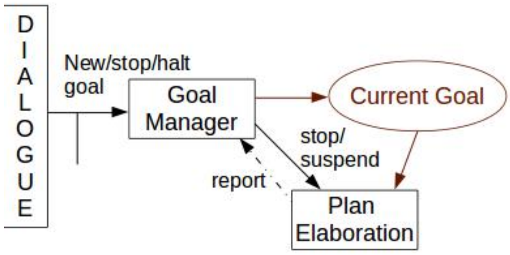


Figure 2.4: Interaction of the Goal Manager with the rest of the supervisor.

activities. This will allow the robot to choose if it should proactively offer its help based on this data and the goal orders it received.

2.2.2 Plan elaboration

Once a goal received from the **Goal Manager**, the **Plan Elaboration** module is in charge of finding a Shared Plan to perform it. To do so, the module is able to call HATP (the Human Aware Task Planner described in Sec. 2.1) to compute a plan and the dialogue module to validate the plan or ask for missing information. One of the contributions of this thesis concerns the elaboration of more flexible Shared Plans where some decisions are left to the execution. This work is in part done in this module and will be developed in Chapter 4.

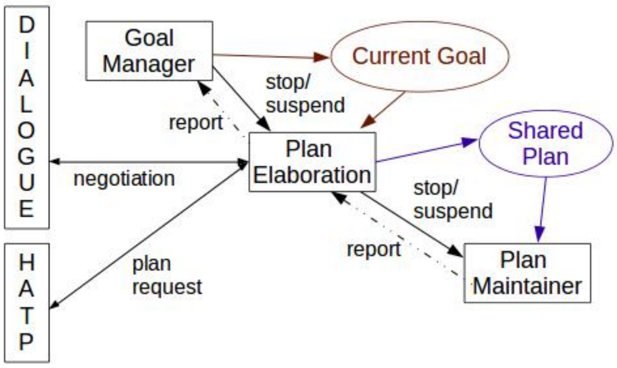


Figure 2.5: Interaction of the Plan Elaboration with the rest of the supervisor.

The computed Shared Plan is then published in order for the **Plan Maintainer** module to deal with it. The *stop* and *suspend* orders received from the **Goal Manager** are transmitted to the **Plan Maintainer** module from which it receives reports concerning the success, failure or need of adaptation of the Shared Plan.

2.2.3 Plan Maintainer

The **Plan Maintainer** module is in charge of monitoring the execution of the Shared Plan based on the current world state and current and past actions. It publishes the list of actions from the Shared Plan which need to be performed at a given moment and the list of actions from the Shared Plan which need to be done later. It also checks the consistency of the plan and reports to the **Plan Elaboration** module in case of failure or unexpected situations.

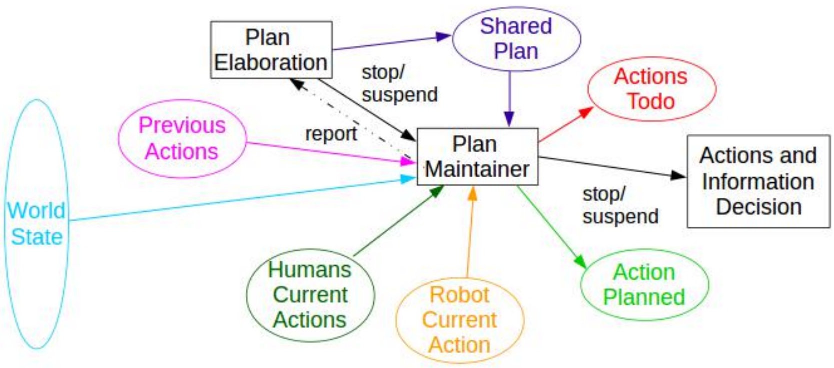


Figure 2.6: Interaction of the Plan Maintainer with the rest of the supervisor.

2.2.4 Human Monitor

The **Human Monitor** module allows to interpret the current world state which contains humans activity information in order to recognize basic humans actions like Pick or Place. This module is, for now, really basic as it is based mainly on distances between humans and objects. However, there is room for improvements by taking into account the context (e.g. the action the agent is supposed to perform) during action recognition or using probabilistic models.

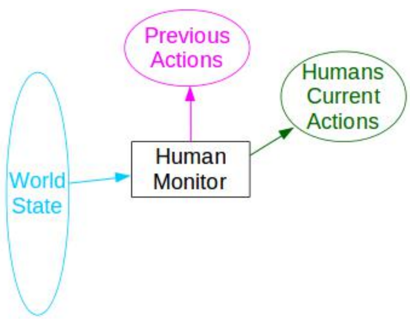


Figure 2.7: Interaction of the Human Monitor with the rest of the supervisor.

2.2.5 Mental State Manager

The **Mental State Manager** estimates the humans mental states concerning the current goal and Shared Plan. It bases its reasoning on all data published by the other supervisor modules and on the world states from all agents point of view given by TOASTER (see Sec. 2.1). This work is one of the thesis contributions and will be developed in Chapter 3. The composition of the estimated mental states will be given in Sec. 2.3.

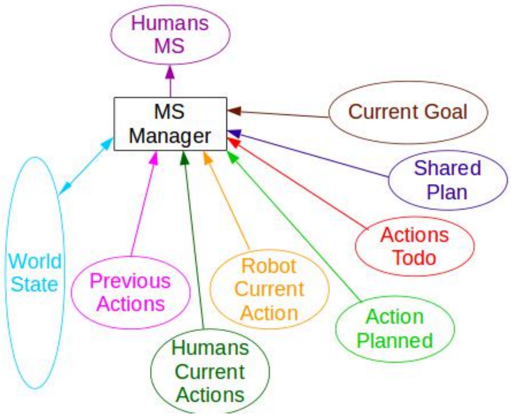


Figure 2.8: Interaction of the Mental State Manager with the rest of the supervisor.

2.2.6 Actions and Information Decision

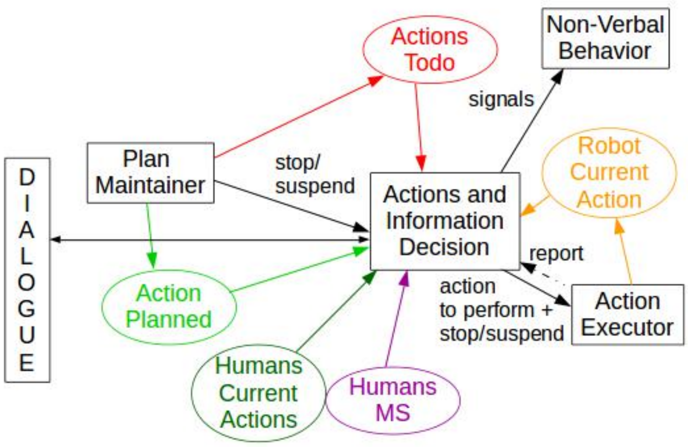


Figure 2.9: Interaction of the Actions and Information Decision module with the rest of the supervisor.

The **Actions and Information Decision** module allows the robot to decide which action to execute and which information to give. Its decisions are based on the lists of current, planned and to execute actions as well as on the humans mental states. The way the robot uses these mental states to give pertinent information to humans is one of the thesis contribution and will be developed in Chapter 3. The decision of which action to execute has also been studied in this thesis and will be developed in Chapter 4.

The **Actions and Information Decision** module sends commands to the **Action Executor** from which it receives reports. It also communicates with the dialogue module and the **Non-Verbal Behavior** module to give the correct information.

2.2.7 Action Executor

The **Action Executor** is in charge of supervising robot actions. It receives actions to execute and *stop* or *suspend* orders from the **Actions and Information Decision** and calls lower level modules to perform the given action in the best possible way.

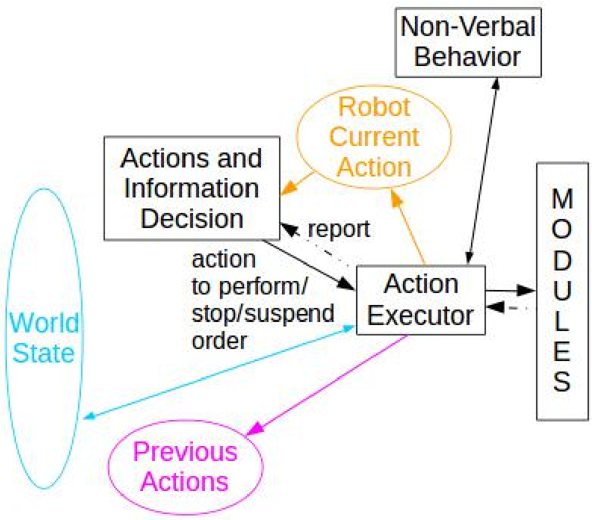


Figure 2.10: Interaction of the Action Executor with the rest of the supervisor.

2.2.8 Non-Verbal Behavior

This module allows to control the non-verbal behavior of the robot. In the current supervisor version, only the robot head behavior is concerned, but other types of non-verbal behaviors can be envisioned. The principles behind this module will be developed in Chapter 6.

The robot head behavior is based on the current robot action, the humans activities and the actions to perform. The module communicates with the dialogue module in order to coordinate and calls lower modules to control the robot.

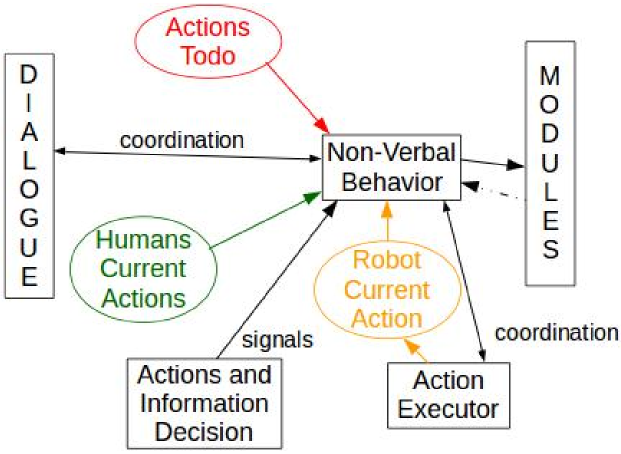


Figure 2.11: Interaction of the Non-Verbal Behavior with the rest of the supervisor.

2.3 Data representation

As seen in the previous section, several types of data are produced and used by the supervisor to take decisions. We will see now how we represent this data and this formalization will be used in the next chapters of the thesis.

The current state of the world from the robot point of view WS is composed of a set of predicates p:

$$p =< entity, attribute, value >$$

For example, the fact that an object is on a table will be represented as

$$< object, isOn, table >$$

.

A goal g is represented as:

$$g =< Name_g, Actors_g, Params_g, Obj_g >$$

where $Name_g$ allows to identify the goal, $Actors_g$ are the agents involved in the goal achievement, $Params_g$ are entities (agents or objects) used to define precisely the goal and Obj_g is a set of predicates representing the objective of the goal. For example, if the robot has for goal to clean the table of the kitchen in collaboration with Bob by removing all items on it, this goal will be represented as:

$$< Clean, < Robot, Bob >, < Kitchen\_table >, << NULL, isOn, Kitchen\_table >>>$$

. Finally, at its end, each goal g is stored and associated with a label noted $label_g$ which can be equal either to DONE or ABORTED.

Then, a Shared Plan SP is represented as:

$$SP =< id_p, A_p, L_p >$$

where id_p is used to identify the plan, A_p are the actions composing the plan and L_p the links representing the order the actions should be executed (causal links).

A link $l \in L_p$ is described as:

$$l = \langle prev_l,\ next_l \rangle$$

where $prev_l$ is the id of the action which needs to be achieved before the action with the id $next_l$ is performed.

The actions composing the plan A_p can be decomposed as:

$$A_p =< A_{prev}, A_{cur}, A_{next}, A_{later} >$$

where A_{prev} are the actions of the plan already executed, A_{cur} the actions currently executed, A_{next} the actions which can be performed according to causal links and actions preconditions and A_{later} the actions to be executed in the future. Each of the set of actions previously introduced can be decomposed as:

$$< A = A^R, A^H, A^X >$$

where A^R are the actions assigned to the robot, A^H the actions assigned to the human and A^X the actions not yet assigned. Indeed, we will see in Chapter 4 that not all actions are assigned to an actor during plan elaboration.

Finally, each action a in A_{prev} is associated with a label noted $label_a$ which can be equal either to DONE, FAILED or ABORTED.

An action a is represented as:

$$a =< id_a,\ Name_a,\ Ag_a,\ Params_a,\ Precs_a,\ Effects_a >$$

Where id_a is the action identifier and $Name_a$ represents its name. $Actors_a$ are the actors of the actions and $Params_a$ a set of parameters (objects or agents) which allows to define precisely the action. $Precs_a$ and $Effects_a$ are sets of predicates representing respectively the action preconditions and effects. For example the action for the robot to place an object on a support will be defined as:

$$< 0, place, Robot, < object, support >, << object, isHoldBy, Robot >>,$$

$$<< object, isOn, support >>>$$

.

The information concerning the state of the task is grouped in what will be called the Task State TS:

$$TS =< g_R, SP, WS >$$

with g_R the current goal of the robot, SP the current Shared Plan and WS the current world state from the robot point of view.

The robot also have a representation of humans mental state. The representation of the mental state $MS(H)$ of a human H is represented as:

$$MS(H) =< g_H, g_R(H), SP(H), WS(H) >$$

where g_H is the goal the robot estimates the human is engaged in, $g_R(H)$ is the goal the robot estimates the human thinks the robot is performing, and $SP(H)$ and $WS(H)$ are the estimation of the Shared Plan and the World State from the human point of view. $SP(H)$ is represented in the same way as the robot Shared Plan.

All these terms are reminded in the Appendix A

Part II

On the use of Shared Plans during Human-Robot Joint Action

Chapter 3

Taking Humans Mental States into account while executing Shared Plans

Contents

3.1 Motivations

When collaborating with humans, it is essential for the robot to not consider humans as obstacles or tools impacting the environment. As humans are social creatures, the robot must take into account their comfort and so, their point of view. Several works already allow robots to estimate humans perspective and beliefs concerning their environment. In order to improve its ability to perform human-robot Joint Action, the robot must be able to take this information into account when taking decision on how to act or what to communicate. Even if several works have been done on how to integrate humans perspective in dialogue or use it to help the understanding of humans behavior, there is still a gap when it comes to use it during Shared Plan execution. This work aims to start filling this gap by extending the robot knowledge on humans mental states to the joint task and using it to better communicate during Shared Plan execution. It has been the subject of a publication at the HRI 2016 conference [Devin 2016a].

3.2 Theory of Mind

3.2.1 Social Sciences literature

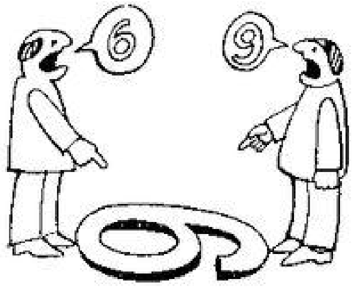
Theory of the Mind (ToM) refers to the ability humans have to recognize and attribute mental states not only to themselves but to other people, to understand that feelings and beliefs we have may be different than the one of the others and to take others mental states into account when taking decisions. ToM has been extensively studied in psychology, particularly in the developmental psychology domain [Baron-Cohen 1985, Premack 1978]. [Verbrugge 2008] defines what is called "order" of ToM:

> "To have a first-order ToM is to assume that someone's beliefs, thoughts and desires influence one's behavior. A first-order thought could be: 'He does not know that his book is on the table'. In second-order ToM it is also recognized that to predict others' behavior, the desires and beliefs that they have of one's self and the predictions of oneself by others must be taken into account. So, for example, you can realize that what someone expects you to do will affect his behavior. For example, '(I know) he does not know that I know his book is on the table' would be part of my second-order ToM. To have a third-order ToM is to assume others to have a second-order ToM, etc."

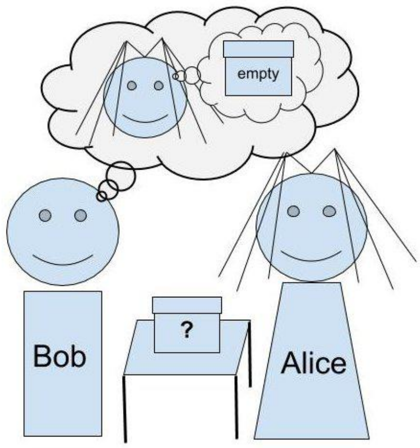
There is an infinite number of orders, however, studies have shown that orders above the second one do not help in cooperative tasks [De Weerd 2014] and those above the third one do not help for competitive games [De Weerd 2014].

ToM includes the notion of perspective taking: the capacity for a person to reason by taking the point of view of someone else. Studied in literature [Tversky 1999, Flavell 1992], perspective taking is crucial during humans interaction and studies have demonstrated that individuals who lack of this ability have difficulties in their daily social interactions [Frick 2014]. Two levels of perspective taking are defined in [Flavell 1977]: perceptual and conceptual perspective taking. Perceptual perspective taking refers to the capacity of a person to understand that others have a different perception of the world (fig D.6(a)). Conceptual perspective taking refers to the capacity of a person to attribute beliefs and feelings to others (fig D.6(b)).

To check if an individual has ToM capacities, several tests have been developed in psychology. One of the most famous is the Sally and Anne test (fin 3.2). This test



(a) Perceptual perspective taking: two individuals can have a different representation of their environment considering their locations.



(b) Conceptual perspective taking: here Bob attributes to Alice a belief concerning the box. He thinks Alice thinks the box is empty.

Figure 3.1: Illustration of perceptual and conceptual perspective taking.

allows to check the capacity of someone to attribute a false-belief to another person and have been reused in robotics to validate robots perspective taking abilities [Hiatt 2010, Milliez 2014].

3.2.2 Robotics background

One of the pioneer work in robotics Theory of Mind is [Scassellati 2002]. Scassellati presents two models from social sciences (Leslie [Leslie 1984] et Baron-Cohen [Baron-Cohen 1997]) and proposes a model on how to implement ToM in robotics. However, the implementation of this model did not go further than perception level.

Then, several works have been done in order to endow robots with perspective taking abilities. Using ACT-R architecture [Anderson 2004], the team of Hiatt and Trafton models mechanisms used during the Sally and Anne test and constructs a model that learns to deal with false belief in order to pass this test [Hiatt 2010]. They extend this work to second-order in [Hiatt 2015] and to spatial reasoning in [Hiatt 2004]. The Sally and Anne test has also been passed in [Milliez 2014] where the robot constructs a semantic representation of the world from its partners point of view. In [Berlin 2006], authors present a way to record different beliefs of other agents and so to have a memory of perspective taking. Finally, [Johnson 2005b] presents a system which computes perspective taking based on forward and inverse visual models.

Perspective taking abilities have been used in robotics for several purposes. It has been used in [Hiatt 2011] to deal with uncertainty in humans behavior and in [Ros 2010b] to solve ambiguous references to an object. One important application of perspective taking is action recognition. [Johnson 2005a] takes the visual point of

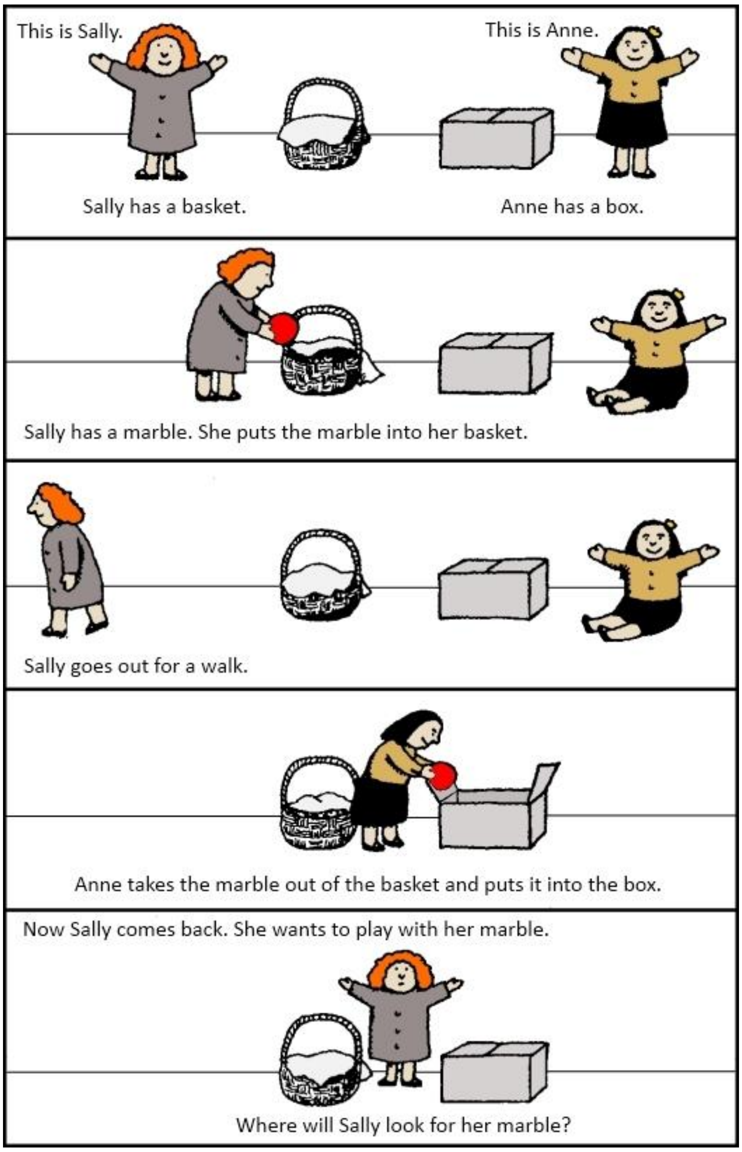


Figure 3.2: The Sally and Anne test: it allows to check the capacity of someone to attribute a false-belief to another person. Illustration from the work of Axel Scheffler.

view of humans to improve action recognition, Dynamic Bayesian Networks (DBN) are used in [Baker 2014] or inverse reinforcement learning in [Nagai 2015]. The human perspective is also used in [Breazeal 2006] to learn a task from a situation that can be ambiguous from the robot point of view and in [Gray 2014] to choose actions with the adequate effects in order to manipulate humans mental models. Finally, [Görür 2017] uses perspective taking to infer humans intention and adapt robot decision.

Concerning Shared Plans, perspective taking can be used to help their elaboration in order to add communication actions to solve divergent beliefs [Warnier 2012]. Then, the human perspective is used to share the plan with a level of details depending of human knowledge [Milliez 2016b]. However, there is no previous work concerning the management of Shared Plans execution taking into account the human point of view.

3.3 Assumptions

The work presented in this chapter concerns the estimation of humans knowledge on the task and its use to help the Shared Plan execution. To do so, we make several assumptions:

Commitment: we do not focus in this work on issues related to commitment. Consequently, we consider here that the joint goal has already been established. We also consider that none of the humans will abort the goal unless he knows that the goal is not achievable any more.

Shared Plan: the focus of this work concerns rather the Shared Plan execution than the Shared Plan elaboration. In the examples presented in Sec. 3.6, the Shared Plan is computed by the robot, however, the processes presented in this chapter hold regardless of the way the robot gets the Shared Plan (e.g. it can be imposed by a human or negotiated through dialogue). This chapter will treat only the issues related to ToM usage in Shared Plan execution, the other aspects of Shared Plan management will be further developed in Chapter 4.

ToM order: this work implements a first-order ToM for the robot (i.e. the robot has knowledge about the human knowledge on the task), the higher orders are not managed for now. It means that the robot has its knowledge concerning the task (0-order) and the estimate of the knowledge of the human concerning the task (1st order) but does not have knowledge concerning what the human knows about what the robot knows about the task (2nd order).

Humans perception: we make the assumption here that a human will see and understand an action of another agent (mainly robot actions) when he is present and looking at the agent. We also assume that when he is present, the human is able to hear and understand the information verbalized by the robot.

Robot capacities: we consider that the robot is able to perform simple high level actions like Pick, Place or Drop. We also assume that the robot is able to ask to a human to perform an action and to inform him about the state of the environment, the goal or an action. The robot is able to detect and localize objects and agents and to recognize simple high level actions performed by a human like Pick, Place or Drop. Let us also note that the ways the robot achieves actions (e.g. human-aware motion planning and execution) and recognizes humans' actions are outside of the scope of this chapter.

Communication: this work consists mainly in finding which information to give to the human and when. We do not focus here on how we give this information (here we use the basic dialogue module described in Chapter 2 but more complex communication mechanisms can be envisioned).

3.4 Estimating Humans Mental States

As stated previously, the goal of this work is to fill the gap between existing perspective taking abilities of the robot and Shared Plan execution. A first step to do so is to extend the knowledge of the robot concerning humans mental states to information concerning the Shared Plan. As seen in Chapter 2, the mental state of a human H will be described as:

$$MS(H) =< g_H, g_R(H), SP(H), WS(H) >$$

where g_H is the goal the robot estimates the human is engaged in, $g_R(H)$ is the goal the robot estimates the human thinks the robot is performing, and $SP(H)$ and $WS(H)$ are the estimation of the Shared Plan and the World State from the human point of view.

The process to estimate the humans mental states will be noted in the following of the thesis as the operator:

$$MS(H) \leftarrow ESTIMATE\_MS(MS(H), TS)$$

with TS the state of the task from the robot point of view as stated in Chapter 2. We will see now how we estimates each of the mental states components. The terms used in the following algorithms and formulas are reminded in Appendix A.

3.4.1 Goal management (g_H and $g_R(H)$ computation)

As stated previously, we do not focus in this work on issues related to goal management. Consequently, the computation of humans mental states concerning goals remains basic. However, a more complex one can be envisioned, for example using intention recognition, with the same representation. As a reminder, a goal is defined as:

$$g =< Name_g, Actors_g, Params_g, Obj_g >$$

As we consider humans automatically engaged in the goal, as soon as the robot starts executing a goal, all actors of the goal are considered to have the same goal:

$$\forall H \in Actors_{g_R},\ g_H = g_R$$

We also make the basic assumption that all humans who see the robot are aware of its goal:

$$\forall H \mid < Robot, isVisibleBy, H > \in WS,\ g_R(H) = g_R$$

For a goal to be considered achieved by an agent (it holds for human mental states as well as the robot mental state), this agent needs to have all the objectives of the goal in its knowledge (it means that according to its knowledge, the desired world state has been reached):

$$\forall H,\ \forall g \mid Obj_g \in WS(H),\ label_g = DONE$$

The robot will consider a goal failed if it can not find any plan to achieve it. Concerning the humans, they can be informed through dialogue by the robot of the failure (or success) of a goal.

3.4.2 Shared Plan management ($SP(H)$ computation)

As a reminder, the representation of the Shared Plan SP from a human H point of view is represented as:

$$SP(H) =< id_p(H), A_p(H), L_p(H) >$$

where $id_p(H)$ is used to identify the plan, $A_p(H)$ are the actions composing the plan and $L_p(H)$ the links representing the order the actions should be executed (causal links).

As we consider in this thesis Shared Plans with action allocation evolving during the execution, we have made the choice 1) to communicate about actions only when it is not implicit and 2) not share the whole plan (more details in Chapter 4). Hence, we consider that the Shared Plan of the human is always the same as the robot one, only the state of the actions composing the plan will change:

$$SP(H) =< id_p, A_p(H), L_p >$$

A link $l \in L_p$ can be described as:

$$l = \langle prev_l,\ next_l \rangle$$

where $prev_l$ is the id of the action which needs to be achieved before the action with the id $next_l$ can be performed.

The actions composing the plan $A_p(H)$ can be decomposed as:

$$A_p(H) =< A_{prev}(H), A_{cur}(H), A_{next}(H), A_{later}(H) >$$

where $A_{prev(H)}$ are the actions of the plan the human thinks are already executed, $A_{cur}(H)$ the actions the human thinks are being executed, $A_{next}(H)$ the actions the human thinks can be performed and $A_{later}(H)$ the actions the human thinks have to be executed in the future.

Each action a in $A_{prev}(H)$ is associated with a label noted $label_a$ which can be equal either to DONE, FAILED or ABORTED.

By default, when a Shared Plan is computed by the robot, all actions are put in $A_{later}(H)$. When the robot performs an action or detects an action execution from a human, it considers the human is aware of the action if he can see the actors of the action:

$$a \in A_{cur} \ \& \ (< Ag_a, isVisibleBy, Human > \in WS \parallel Human \in Ag_a) \\ \Rightarrow a \in A_{cur}(H)$$

Likewise, at the end of the execution, the action goes in $A_{prev}(H)$ with the label corresponding to the success or the failure of the action if the human performs or has seen the actors of the actions at the end of the action.

We also consider that a human can infer that an action has been done if he knows that the action was in progress or needed to be done and he can see the effects of the action:

$$(a \in A_{cur}(H) \parallel a \in A_{next}(H)) \ \& \ Effects_a \in WS(H) \\ \Rightarrow a \in A_{prev}(H) \ \& \ label_a(H) = DONE$$

Likewise, we consider that if a human knows that an action was in progress and can see the actors of the action while there are not performing the action any more, he considers the action DONE:

$$a \in A_{next}(H) \ \& \ a \in A_{prev} \ \& \ < Ag_a, isVisibleBy, Human > \in WS \\ \Rightarrow a \in A_{prev}(H) \ \& \ label_a(H) = DONE$$

Finally, the actions are set in $A_{next}(H)$ considering causal links and preconditions:

$$a \in A_{next}(H) \Leftrightarrow Precs_a \in WS(H) \ \& \ (\forall l \in L_p \\ \mid next_l = id_a, \exists \ ap \in A_{prev}(H) \mid (id_{ap} = prev_l \ \& \ label_{ap}(H) = DONE))$$

The robot can also inform a human about the state of an action, in which case the given information will be added to the human mental state.

3.4.3 World State management ($WS(H)$ computation)

We saw that the perspective taking abilities of the robot allow it to estimate the human perception of his environment [Milliez 2014]. However, this previous work only concerns information about the environment which is perceivable. Indeed, for this work we will consider two kinds of predicates to describe the state of the environment:

— **Observable predicates:** they concern what the agent can observe about the world state. These predicates mainly represent the affordances of all agents (e.g. isVisibleBy, isReachableBy) and the relations between objects (e.g. isOn, isIn) visible to them. They are computed continuously by the Situation Assessment module (TOASTER) from the robot and humans point of views based on geometric computations and perspective taking algorithms.

— **Non-observable predicates:** they concern information that the agent can not observe (e.g. the fact that an opaque bottle is empty or full). These predicates are not managed by TOASTER which reasons only on what is visible by the agents. We consider two ways for an agent to be aware of a non-observable predicate. It can perform or see another agent performing an action which has this predicate in its side effects (e.g. every agent who see another agent filling an opaque bottle know that the bottle is full):

$$\forall a \in A_{prev} \mid label_a = DONE \Rightarrow Effects_a \rightarrow WS$$

(likewise with $A_{prev}(H)$ and $WS(H)$). A agent can also be aware of a non-observable predicate if he is informed of it by another agent.

3.5 Mental States for Shared Plans execution

We saw in the previous section how we estimate humans mental states concerning the shared task. We will see now how we use them to communicate during the Shared Plan execution. Indeed, when two humans share a plan, they usually do not communicate all along the plan execution. Only the *meshing subplans* of the plan need to be shared [Bratman 1993]. Consequently, the robot should inform humans about elements of the shared plan only when it considers that the divergent belief might have an impact on the joint activity in order to not be intrusive by giving them information which they do not need or which they can observe or infer by themselves. The process of monitoring the divergent beliefs and solving them if needed which will be described in this section will be noted in the following of the thesis as:

$$SOLVE\_DB(MS(H), TS)$$

3.5.1 Weak achievement goal

If we follow the definition of *weak achievement goal* in [Cohen 1991], if the robot knows that the current goal has been achieved or is not possible anymore, it has to inform its partners. Accordingly, we consider that, when, in the robot knowledge, the label of a goal is DONE (resp. ABORTED) and the robot estimates that a human does not consider it DONE (resp. ABORTED), the robot informs him about the achievement (resp. abandoning) of the goal (if the agent is not here or is busy with something else, the robot will do it as soon as the agent is available again).

Algorithm 1 Weak achievement goal

if $\exists\ g \mid (label_g = DONE\ \&\ label_g(H) \neq DONE) \parallel (label_g = ABORTED\ \&\ label_g(H) \neq ABORTED)$ ▷ *There is a divergent belief to solve*
 then
  Inform(g) ▷ *The robot informs the human about the state of the goal*
end if

3.5.2 Before humans action

A divergent belief of a human partner can be an issue when it is related to an action that he has to perform. To avoid that a human misses information to execute his part of the Shared Plan, each time the robot estimates that a human has to perform an action (action in A^H_{next}) it checks if the human is aware that he has to and can perform the action (the action should also be in $A^H_{next}(H)$). In order not to give too much information at the same time, and as not yet allocated actions can also be performed by the robot, the robot checks for actions in $A^X_{next}(H)$ only if the human does not have any other actions to perform. If there is a divergent belief, there are two possible reasons:

— The human misses information about previous achieved actions to know that his action has to be performed now according to the plan. The robot checks the label of all actions linked to the first one with the plan links. If it finds one with a label different of DONE in the estimation of the human knowledge it informs about its achievement.
— The human misses information about the world state to know that his action is possible. In such case, the robot looks into the preconditions of the actions and informs the human about all those the human is not aware of.

The robot first looks for missing actions before looking for missing preconditions. Indeed, informing about an action can also give information about some missing preconditions. The given algorithm to solve this kind of divergent belief is summarized in Alg. 2.

Algorithm 2 Checking humans actions

if $(\exists\ action \in A^H_{next} \mid action \notin A^H_{next}(H)) \parallel (A^H_{next} = \emptyset\ \&\ \exists\ action \in A^X_{next} \mid action \notin A^X_{next}(H))$ **then**
▷ *There is a divergent belief to solve*
  if $\exists\ a \in A_p \mid (\exists\ l \in L_p \mid next_l = action\ \&\ prev_l = a)\ \&\ (a \notin A_{prev}(H) \parallel label_a(H) \neq DONE))$ **then**
    Inform(a) ▷ *The robot says to the human that the action is done*
  end if
  if $\exists p \in Precs_{action} \mid p \notin WS(H))$ **then**
    Inform(p) ▷ *The robot informs about the missing precondition*
  end if
end if

3.5.3 Preventing mistakes

A divergent belief of a human partner can also be an issue if it leads him to perform an action that should not be performed now according to the plan. To prevent this, for each action that the robot estimates the human thinks he has to execute (action in $A^{H}_{next}(H)$), the robot checks if the action really needs to be performed (the action should also be in A^{H}_{next}). In the same way, the robot also checks the actions not yet allocated as the human can perform them (action in $A^{X}_{next}(H)$ which is not in A^{X}_{next}). If there is a divergent belief, the robot corrects the human divergent belief by two different ways:

— The human can think that a previous action has been achieved successfully while it is not the case (e.g. he saw the beginning of an action of the robot and though that the action succeeded) leading him to think he has to perform another action. The robot looks in all actions linked to the first one by the plan links and informs about their state if it is different in the estimation of the human knowledge and in the robot knowledge.

— The human can have a divergent belief concerning the world state that leads him to think that his action is possible while it is not the case. The robot looks into the preconditions of the action and informs about divergent beliefs.

The given algorithm to solve this kind of divergent belief is summarized in Alg. 3.

Algorithm 3 Preventing mistakes

if $\exists\ action \in A^{H}_{next}(H)\ \cup\ A^{X}_{next}(H) \mid action \notin \{A^{H}_{next}\ \cup\ A^{X}_{next}\}$ **then**
▷ *There is a divergent belief to solve*
  if $\exists\ a\ \in\ A_p\ \mid\ (\exists\ l\ \in\ L_p\ \mid\ next_l\ =\ action\ \&\ prev_l\ =\ a)\ \&\ (label_a(H)\ = DONE)\ \&\ (a \notin A_{prev}\ \|\ label_a \neq DONE))$ **then**
    Inform(a) ▷ *The robot informs about the state of the action*
  end if
  if $\exists p \in Precs_{action} \mid p \in WS(H)\ \&\ p \notin WS)$ **then**
    Inform(p) ▷ *The robot informs about the wrong precondition*
  end if
end if

3.5.4 Signal robot actions

When the robot is about to perform an action, it checks if it estimates that the humans are aware that it will act (the action should also in $A^{R}_{next}(H)$). If it is not the case, the robot signals its action before performing it (Alg. 4).

Algorithm 4 Signal robot actions

if $\exists\ action \in A^{R}_{next} \mid action \notin A^{R}_{next}(H)$ ▷ *There is a divergent belief to solve*
then
  Signal($action$) ▷ *The robot signals its action*
end if

3.5.5 Inaction and uncertainty

Even if the robot estimates that the human is aware that he has to act (there is action(s) in $A^H_{next}(H)$), it is possible that the human still does not perform his action(s). If the human is already busy (there is an action in A^H_{cur}) or if he is not currently engaged in the task, the robot waits for the human to be available. If the human is not considered busy by the robot, the robot first considers that its estimation of the human mental state can be wrong, and that, in reality, the human is not aware that he should act. Consequently, the robot asks the human specifically to perform the action. If the human still does not act while the action has been asked, the robot considers the action failed, aborts the current plan and tries to find an alternative plan excluding that action.

3.6 Results

3.6.1 Tasks

In order to evaluate the benefits of our method during human-robot interaction, we will use two different tasks. In a first time, we will show an illustrative example based on one possible scenario of the first task. In a second time, we will run simulations on the two tasks in order to get objective measurement of the performance of the system in simulation. The system will be evaluated in a real situation later in Chapter 5.

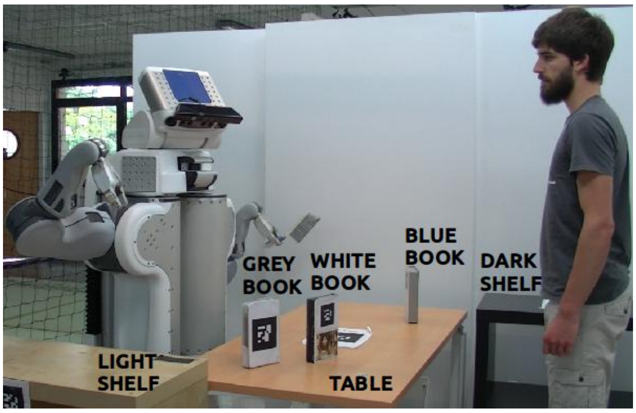


Figure 3.3: Initial situation of the "Clean the table" scenario.

"Clean the table" scenario In this example, a PR2 robot and a human have to clean a table together. To do so, they need to remove all items from this table, sweep it, and re-place all previous items. The initial world state is the one in Fig. D.7. We consider that the grey book is reachable only by the robot, the blue book only by the human and the white book by both agents. The human and the robot have the ability to pick objects and place them into another support. Only the robot has the capacity to sweep the table. The initial plan produced to achieve the goal is shown in Fig. D.8.

"Inventory" scenario In this example, a human and a PR2 robot have to make an inventory together. At the beginning of the task, both agents have colored objects near them as well as a colored box (initial world state in Fig. 3.4). These colored objects need to be scanned and then, stored in the box of the same color. To do so, both agents can pick objects and place them in the table in a way reachable by the other agent. They also both have the ability to drop objects in the box near them. Finally, only the robot can scan an object, it consists of orienting its head and turning on a red light in the direction of a reachable object.

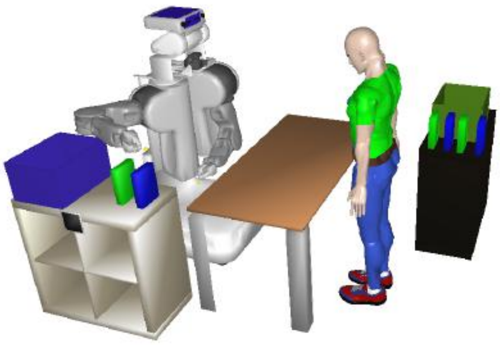


Figure 3.4: Initial situation of the "Inventory" scenario.

3.6.2 Illustrating scenario

We will first illustrate the benefits of the work presented in this chapter with an example. This example is based on the "Clean the table" task presented previously. At the beginning of the interaction the robot computes the plan Fig. D.8.

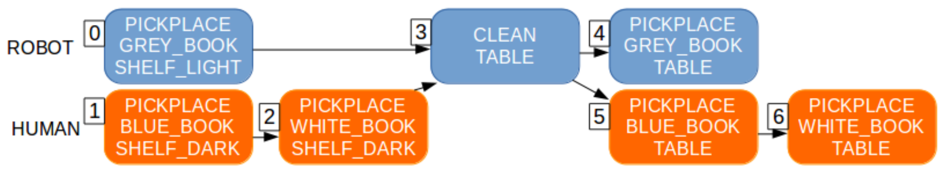


Figure 3.5: Initial plan of the "Clean the table" scenario.

The robot starts to pick and place the grey book on the light-colored shelf. The human picks and places the blue book on the dark-colored shelf then leaves (Fig. D.10(a)).

Robot				Human			
A_{prev}	A_{cur}	A_{next}	A_{ready}	A_{prev}	A_{cur}	A_{next}	A_{ready}
1	0	2	3, 4, 5, 6	1	0	2	3, 4, 5, 6

Table 3.1: Knowledge of the robot and estimation of the human knowledge after the human has left. The numbers represent the actions id as stated in the plan Fig. D.8.

The robot ends its action. At this point, the only possible action of the plan has to be done by the human. The robot waits a given amount of time for the human and then, as the human does not come back (and so does not execute his action), aborts the current plan and computes a new one where it removes the last book itself (Fig. D.9).

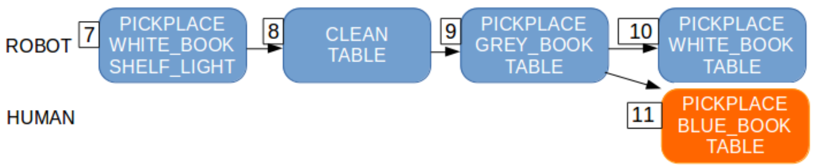


Figure 3.6: Second plan of the "Clean the table" scenario.

The robot picks and places the white book on the light-colored shelf and sweeps the table (Fig. D.10(b)).

Robot				Human			
A_{prev}	A_{cur}	A_{next}	A_{ready}	A_{prev}	A_{cur}	A_{next}	A_{ready}
0, 1, 7, 8		9, 11	10	1	0	7	8, 9, 10, 11

Table 3.2: Knowledge of the robot and estimation of the human knowledge after the robot swept the table. The numbers represent the actions id as stated in the plan Fig. D.9.

The human comes back at this time (Fig. D.10(c)). As he can see that the grey book is on the shelf near the robot, the robot infers that he infers that the robot has achieved the action it was performing when the human left. Moreover, as the human can see that the white book is on the shelf near the robot, the robot infers that he infers that the robot moved the book. However, the human can not observe that the table has been swept by the robot (we consider here that the effects of the *sweep* action are not observable). Consequently, the human does not know that he can put back the book he removed.

Robot				Human			
A_{prev}	A_{cur}	A_{next}	A_{ready}	A_{prev}	A_{cur}	A_{next}	A_{ready}
0, 1, 7, 8		9, 11	10	0, 1, 7		8	9, 10, 11

Table 3.3: Knowledge of the robot and estimation of the human knowledge when the human comes back. The numbers represent the actions id as stated in the plan Fig. D.9.

As the robot estimates that the human does not know that he has to put back the book he removed, it uses its knowledge on the plan to infer that it is because the human does not know that the table has been swept. So, the robot informs the human about this (by verbalization).

Robot				Human			
A_{prev}	A_{cur}	A_{next}	A_{ready}	A_{prev}	A_{cur}	A_{next}	A_{ready}
0, 1, 7, 8		9, 11	10	0, 1, 7, 8		9, 11	10

Table 3.4: Knowledge of the robot and estimation of the human knowledge after the robot informed the human. The numbers represent the actions id as stated in the plan Fig. D.9.

The human has now all the information he needs to finish the task. The robot and him both perform their last actions and so achieve the task (Fig. D.10(d)).

3.6.3 Quantitative results

We will now evaluate the benefits of the presented work in simulation in the two tasks described previously. Results in real situations as well as more simulation results with the whole system developed in the thesis can be found in Chapter 5). The results here only concern the use of mental states during Shared Plan execution.

When the interaction starts, we consider that the joint goal is already established and that a first Shared Plan has been computed by the robot. The robot executes the plan and the simulated human executes the actions planned for him. We randomly sample a time when the human leaves the scene and another time when the human comes back. While absent, the human does not execute actions and cannot see anything nor communicate.

One objective of our contribution is to reduce unnecessary communication from the robot during the execution of a Shared Plan aiming at a more friendly and less intrusive behavior of the robot. Consequently, in order to evaluate our system, we have chosen to measure the amount of information shared by the robot during a Shared Plan execution. During the interaction, we logged the number of facts (information chunks) given by the robot to the human. An information concerns either a change in the environment or the state of a previous action.

We compared our system (called *ToM system*) to:

(a) The human leaves after removing his first book from the table.

(b) The robot removes the last book and sweep the table.

(c) The human comes back.

(d) The human and the robot perform their last actions and achieve the task.

Figure 3.7: Illustrative "Clean the table" scenario.

— a system which informs about each action missed by the human (called *Missed system*).

— a system which informs about each action performed by the robot even if the human sees it (called *Performed system*).

The obtained results in 100 runs are given in Table 3.5.

We can see that our system allows to reduce significantly the amount of information given by the robot. In the "Clean the table" scenario, depending on when the human leaves, the robot might change the initial plan and take care of the book reachable by both agents instead of the human. This explains why the average number for the *Performed system* is higher than the number of actions initially planned for the robot: the robot performs more actions in the new plan. In this scenario, our system allows not to communicate about missed *pickandplace* actions as the human can infer them by looking at the objects placements. However, the robot will inform the human if he missed the fact that the robot has swept the table as it is not observable and it is a necessary information for the human to know before he can put back objects on the table.

In the inventory scenario, as all objects and boxes are reachable only by one agent, the robot does not need to change the plan when the human leaves. This explains the fact that the standard deviation is null for the *Performed system*: the number of actions performed by the robot never changes and there is no change in

Scenario System	Clean the table		Inventory	
	Average	Std Dev	Average	Std Dev
ToM	0.94	0.24	0.41	0.48
Missed	2.14	0.87	2.61	1.36
Performed	3.72	0.96	10.0	0.0

Table 3.5: Number of information given by the robot during the two presented scenarios for the three systems (*TOM*, *Missed* and *Performed*).

the plan. In this scenario, the *pickanddrop* and *scan* actions have non-observable effects (the human can not see an object in a box). However, we can see that our system still verbalizes less information than the *Missed system*: the robot communicates only the information which the human really needs (as the fact that an object the human should drop in a box has been scanned) and does not give information which are not linked to the human part of the plan (as the fact that the robot scanned an object it have to drop in its box or that the robot dropped an object).

3.7 Conclusion

In this chapter we showed how we extended the robot estimation of humans mental states (which initially concerned only the environment) to the state of the task and more specifically of the Shared Plan. Then, we showed how we use these mental states to better communicate during Shared Plan execution.

The benefits of this work have been demonstrated with an illustrative example and simulation results. These results show that the proposed system allows to reduce communication by removing useless information given by the robot.

The next chapter will give more details about Shared Plan elaboration and management, more particularly concerning actions allocation. Then, we will show in Chapter 5) more complete simulation results with the whole system as well as results with the system running in a real situation.

Chapter 4

When to take decisions during Shared Plans elaboration and execution

Contents

4.1 Motivation

When performing a Joint Action and more particularly when executing Shared Plans, several choices have to be made. Some of them are implicit, while others require a negotiation or an adaptation between the Joint Action participants. To be a good partner when performing Joint Action with humans, the robot should be able to identify which decisions are implicit and correctly communicate about the other ones. Indeed, a robot which communicates about each detail of a Shared Plan would quickly become too "chatty" while a robot which does not communicate can be confusing.

Let's take for example a robot helping a human to build a flat-pack table: the legs of the table need to be assembled with a hammer, the tray with a screwdriver and finally someone needs to put the tray on the legs. The robot is equipped with several tools including a screwdriver but no hammer and the human has only one hammer. The human and the robot are both able to put the tray on the legs. It is common sense that the robot should assemble the tray while the human assemble the legs. However, a decision needs to be taken concerning who will put the tray on the legs. In this scenario, we would like the robot to assemble the tray without asking the human and negotiate or adapt its behavior to put the tray at the last moment.

The work presented in this chapter consists in finding which decisions are implicit or not, when the decisions should be taken and how to take them. We identify three types of decisions to be taken during Shared Plans elaboration and execution:

— **Which action to perform in which order:** this is one of the biggest concern during Shared Plan elaboration. We do not focus our work in this challenge. Indeed, we use HATP, a human-aware HTN planner which has been demonstrated to be well suited to human-robot joint action [Lallement 2014], to deal with this issue.

— **Who will perform which action:** sometimes this decision can be trivial when only one agent is able to perform an action. However, in other situations, the robot should be able to decide who will perform an action by negotiating or adapting its behavior to the human one.

— **With which object:** for practical reason, the robot reasons on objects by attributing them a unique id. However, for the purpose of an action, two objects can be semantically identical. When there is a choice in which object to use for an action, the robot should be able to adapt to the human behavior in order to avoid potential conflicts.

This work has been the subject of a publication at the ICSR 2017 conference [Devin 2017].

4.2 Background

When the robot needs to achieve a joint goal, several works allow it to compute plans which take into account the human ([Cirillo 2010, Lallement 2014]). They allow the robot to reduce resource conflicts [Chakraborti 2016], take divergent beliefs into account ([Warnier 2012, Talamadupula 2014]) or promote stigmergic collaboration for agents in co-habitation [Chakraborti 2015].

The relevance of using a Shared Plan in human-robot interaction has been studied by [Lallée 2013]. They suggest that the joint plan should be fully communicated in order to sustain effective collaboration. Moreover, in [Gombolay 2015], it is shown that subjects prefer letting the robot plan when the task is too complex, prioritizing efficiency. In more simple tasks, a robot proactively helping the human is preferred to one waiting before proposing help [Baraglia 2016].

If the robot decides to share the plan, several studies have been reported on how to communicate about the plan. Some researchers studied how a system could acquire knowledge on plan decomposition from a user [Mohseni-Kabir 2015] and how dialog can be used to teach new collaborative plans to the robot and to modify these plans [Petit 2013]. In [Sorce 2015], the system is able to learn a plan from a user and transmit it to another user. [Allen 2002] presents a computer agent able to construct a plan in collaboration with a user. Finally, in [Milliez 2016b], Milliez et al. present a system where the robot shares the plan with a level of details which depends on the expertise of the user. In our work, we try to get rid of the entire shared plan verbalization by taking the right decision at the right time in order to come up with a robot which communicates at the right time.

Several contributions have been done to allow more adaptability during human-robot Shared Plan execution. [Chien 2000] proposes a method to plan only a few steps in advance and then plan the actions further in an iterative way. This allows the plan to incorporate execution feedback such as early or late execution of actions and over-use or under-use of resources. Chaski, a task-level executive that uses insights from human-human teaming to make human-robot teaming more natural and fluid, is presented in [Shah 2011]. The system chooses and schedules the robot's actions by taking into account the human partner and acts to minimize the human's idle time. A system which mixes plan recognition and adaptation is described in [Levine 2014]. It computes all possibilities for the plan and chooses an action based on the choice of the human and causal links. [Hoffman 2007] proposes an adaptive action selection mechanism for a robotic teammate, making anticipatory decisions based on the confidence of their validity and their relative risk. [Karpas 2015] presents Pike, an online executive that unifies intent recognition and plan adaptation for temporally flexible plans with choice. Finally, the work presented here is based on SHARY [Clodic 2009] which was extended in [Fiore 2014], a supervisor allowing to execute human-aware shared plans taking into account joint actions aspects like reactive action execution.

In the cooperative multi-robot literature, task allocation and cooperative activity achievement has been thoroughly investigated [Gerkey 2004]. Auction has been used very successfully for distributed multi-robot in various contexts ([Gerkey 2002, Botelho 1999]). Several works on teamwork and cooperative task achievement take into account explicit constraints to facilitate the activity of the other robots or agents activity. [Tambe 1997] presents STEAM, an architecture based on the *joint intention* theory. It integrates concepts such as team synchronisation and monitoring of joint intention in order to improve flexibility and reusability. A plan manager which provides the services needed to build and execute plans in a multirobot context is presented in [Joyeux 2009]. It provides tools for safe concurrent execution and modification of plans, and handles distributed plan supervision without permanent robot-to-robot communication. While these contributions have inspired work on human-robot collaboration, it is however important to exhibit some differences between the two fields. Indeed, the human and the robot are not equal in any aspect. The robot is here to help the human and facilitate his activity.

In AI, the goal reasoning domains deals with some problems similar to Shared Plan management [Molineaux 2010, Roberts 2016]. The role of goal reasoning is to survey the current goals of a robot, check that they remain feasible and relevant and establish new goals if needed. Moreover, part of the goal reasoning function is sometimes linked to the plan management as it is in charge of deciding when and how to generate plans (but it is not producing the plan) and checking for unexpected events.

4.3 Assumptions

The work presented in this chapter deals with the needed decision during Shared Plan elaboration and execution. The idea here is to focus on decisions concerning the action allocation and instantiation. To do so, we make several assumptions:

Single human: the work presented in this chapter has been designed for a robot interacting with a single human. However, all the data structures and main principles are compatible with multi-humans set-up.

Commitment: we do not focus in this work on issues related to commitment. Consequently, we consider here that the joint goal has already been established. We also consider that the human will not abort the goal unless he knows or infers that the goal is not achievable any more.

Shared Plan: we put the focus here on the issues related to action allocation and instantiation. In order to decide which action to execute in which order, we use HATP, a human-aware HTN planner which has been demonstrated to be well suited to human-robot Joint Action [Lallement 2014]. We are focusing in this work about medium complexity Shared Plans where the human might want to decide for his own actions.

Humans perception: we make the assumption here that a human will see and understand an action of the robot when he is present and looking at it. We also assume that when he is present, the human is able to hear and understand the information verbalized by the robot.

Robot capacities: we consider that the robot is able to perform simple high level actions like Pick, Place or Drop. We also assume that the robot is able to ask to the human if he wants to perform an action and to understand a basic answer (yes/no type). The robot is able to detect and localize objects and agents and to recognize simple high level actions performed by the human like Pick, Place or Drop. Let us also note that the ways the robot achieves actions (e.g. human-aware motion planning and execution) and recognizes human's actions are outside of the scope of this chapter.

Communication: the focus of this work is more on "what to communicate" rather than on "how to communicate". Here we use the basic dialogue module described in Chapter 2 to communicate with the human but more complex communication mechanisms can be envisioned.

4.4 Main principles

We will present in this section the main principles we use for Shared Plans management. Three algorithms are used to allow the robot to elaborate and execute Shared Plans. They interact through the Shared Plan data *SP* and several signals (noted S_X where X is the name of the signal). These three algorithm run constantly and in parallel.They allow respectively to maintain the state of the shared plan, to choose actions for the robot, and to monitor the human.

Alg. 5 allows the robot to elaborate a Shared Plan when needed, to maintain the current Shared Plan and to manage the human mental states. [1]

When the robot has a goal g_R to achieve and no current plan or when a signal is received to compute a new plan ($S\_needReplan$), the robot computes a Shared Plan to perform the goal based on the current world state WS (see Sec. 4.5):

$$SP \leftarrow PLAN(g_R, WS)$$

When an action a from the plan is performed by an agent a signal is received ($S\_needUpdate$) and the supervisor updates the plan (see Sec. 4.6.1):

$$SP \leftarrow UPDATE\_PLAN(SP, a)$$

When an action from A^X has been allocated ($S\_actionAllocated$), the robot looks for the consequences of this allocation in the plan (see Sec. 4.6.3):

$$SP \leftarrow EVALUATE\_PLAN(SP, WS)$$

The robot also constantly checks if the goal is reached (the objectives of the goal are in the current World State). Finally, each time a change occurs in $WS(H)$ or TS, the robot estimates the human mental states as described in the previous chapter (see Chapter 3):

$$MS(H) \leftarrow ESTIMATE\_MS(MS(H), TS)$$

If there is a conflict between the knowledge of the robot and the human mental state, the robot tries to solve it (see Chapter 3):

$$SOLVE\_DB(MS(H), TS)$$

1. The terms used in the algorithms are reminded in Appendix A.

Algorithm 5 Shared Plan management

while g_R **do**
  if $!SP \parallel S\_needReplan$ **then**
    $SP \leftarrow PLAN(g_R, WS)$
  end if
  if S_needUpdate **then**
    $SP \leftarrow UPDATE\_PLAN(SP, WS)$
  end if
  if S_actionAllocated **then**
    $SP \leftarrow EVALUATE\_PLAN(SP, WS)$
  end if
  if $Obj_g \in WS$ ▷ *The goal is achieved*
  then
    $g_R \leftarrow \emptyset$
    $SP \leftarrow \emptyset$
  end if
  if $< Human, isPresent, true > \in WS$ & $(WS(H) \neq WS(H)_{t-1} \parallel TS \neq TS_{t-1})$ **then**
    $MS(H) \leftarrow ESTIMATE\_MS(MS(H), TS)$
    if $MS(H) \neq TS$ ▷ *Divergent belief*
    then
      $SOLVE\_DB(MS(H), TS)$
    end if
  end if
end while

Running in parallel with the first algorithm, Alg. 6 allows the robot to decide when to act and which action to perform.

When the robot has a Shared Plan SP, it looks for the actions of this plan which need to and can be executed (see Sec. 4.6.1):

$$A_{next} \leftarrow GET\_NEXT\_ACTIONS(SP, WS)$$

If there is no action in A_{next} nor in progress, it means that the plan is blocked, so the robot looks for another Shared Plan. If there are actions to do, the robot looks if there is an action it can execute (actions allocated to it or not allocated yet). If there is none, the robot waits for the human to perform an action (A_{next} contains only actions from A^H_{next}):

$$actionExecuted \leftarrow WAIT\_ACTION(A^H_{next}, t)$$

If after a time t, the human did not execute any action, the robot looks for another plan. If there are actions the robot can execute, the robot selects an action a (see Sec. 4.6.2):

$$a \leftarrow SELECT\_ACTION\_TODO(A_{next})$$

If the selected action is not yet allocated, the robot first tries to allocate it (see Sec. 4.6.3):

$$actor \leftarrow ALLOCATE\_ACTION(SP, a, WS, Prefs)$$

If the action is allocated to the robot (after selection or allocation), the robot executes it:

$$success \leftarrow EXECUTE(a)$$

It will first instantiate the action if needed and then launch its execution (see Sec. 4.6.4). If the action succeeds, the robot updates the plan, else it looks for another plan.

Algorithm 6 Robot action decision

```
while SP do
  A_next ← GET_NEXT_ACTIONS(SP, WS)
  if A_next = ∅                                  ▷ No more feasible actions
   then
    S_needReplan
  else if {A^R_next ∪ A^X_next} = ∅              ▷ No actions for the robot
   then
    actionExecuted ← WAIT_ACTION(A^H_next, t)
    if !actionExecuted then
      S_needReplan
    end if
  else
    a ← SELECT_ACTION_TODO(A_next)
    if a ∈ A^X_next then
      actor ← ALLOCATE_ACTION(SP, a, WS, Prefs)
      S_actionAllocated
    end if
    if a ∈ A^R_next || (a ∈ A^X_next & actor = robot) then
      success ← EXECUTE(a)
      if success then
        S_needUpdate
      else
        S_needReplan
      end if
    else if (a ∈ A^X_next & actor = human) then
      a → A^H_next
    end if
  end if
end while
```

In parallel to the other execution loops, the robot is constantly monitoring human activities (Alg. 7).

When the human performs an action, the robot first looks if the action is conflicting to the one it is performing (for example, if the human picks an object the robot was going to pick), and if it is the case, the robot stops its actions. Then, if the human performs an expected action with respect to the plan and already allocated to him, the robot updates the plan accordingly. If the action executed by the human is expected with respect to the plan but was not yet allocated, the robot looks for the consequences of this actions in the plan (see Sec. 4.6.3). If the human performs an unexpected action with respect to the plan (action not in A_{next} or in A^R_{next}), the robot waits the end of the human action and then looks for a new plan from the new situation induced by the human action.

Algorithm 7 Human monitoring

```
while < Human, isPresent, true >∈ WS do
  if ∃ a ∈ A^H_cur then
    if a ∈ A^R_cur then
      S_stop
    end if
    WAIT_END_ACTION(a)
    if a ∈ A^H_next then
      S_needUpdate
    else if a ∈ A^X_now then
      S_actionAllocated
    else
      S_needReplan                                ▷ Unexpected action
    end if
  end if
end while
```

The next sections will define more precisely the operators we just defined.

4.5 Shared Plans elaboration

The first step of this work is to be able to compute Shared Plans that are flexible enough to postpone part of the decisions to execution time. This step correspond to the operator:

$$PLAN(g_R, WS)$$

As stated before, the human-aware HTN task planner HATP is used in this work to compute Shared Plans taking into account a number of social rules for both the robot and its human partner [Lallement 2014]. However, in order to obtain flexible plans with HATP, a number of issues have to be considered.

First, when HATP returns a plan, it returns only one, which is assumed to be the best plan it has found given the situation and the associated costs. However, this plan is not always the only one possible (even at constant cost or computing time). Indeed, in such case, HATP makes some choices that could be preferably done on-line. For example, it can happen that one action can be done by several agents at the same cost. In a collaborative setting and more particularly when the human is concerned, it could be interesting to let the agents decide at execution time (or whenever it is interesting) who will do what. To handle this, we have adapted HATP by inserting what we call the *X agent*. The capabilities of the *X agent* correspond to the intersection of the capabilities of the human and the robot with a lower cost. Consequently, it will be chosen by the planner instead of the human or the robot whenever it is possible. If HATP returns a plan containing an action to be done by the *X agent*, it means that this action could either be performed by the human or the robot. The decision concerning who will finally do this action is postponed. We will see in Sec. 4.6.3 how *X agent* actions will be finally allocated to the human or the robot.

As inputs to a planner such as HATP, we give a set of objects that are present in the environment and on which it will be able to apply its operators. Basically, each object is tagged and is unique. That means that if we have the same object twice, they will be uniquely tagged (e.g. two identical red cubes will be tagged as RED_CUBE_1 and RED_CUBE_2). When two similar objects can be used in a same way during a task, the planner will choose either one or the other. In a collaborative setting, it could be counter intuitive since even if there is no distinction between the two objects at planning time, there can be one during execution. To handle this, we have adapted HATP by inserting the notion of *similar* objects which aims to group interchangeable objects under a common name: two *similar* objects will have the same role in the task.

Finally, rather than in other contexts where we used HATP, we do not consider here that an agent is only capable to perform one action at a time. This allows the human to choose the order of his action when there is no impact in the global plan.

4.6 Shared Plans execution

We will now see in more detail how the robot executes the flexible Shared Plans obtained.

4.6.1 Plan maintaining

First, the robot needs to be able to follow the Shared Plan execution and to determine which actions need to be executed and which actions need to be left for later.

As said before, the actions composing a Shared Plan can be decomposed as:

$$A_p = < A_{prev}, A_{cur}, A_{next}, A_{later} >$$

By default, when a Shared Plan is computed by the robot, all actions are put in A_{later}. When the robot performs an action or detects an action execution from a human, the executed action goes in A_{cur} and, at the end of the execution, the action goes in A_{prev} with a label equal either to:

— DONE if the execution has been successful,
— FAILED if the execution has not been successful,
— ABORTED if the robot had to stop the execution for an external reason.

An action will be put in A_{next} if all previous actions in the plan are DONE (based on causal links) and its preconditions are checked:

$$a \in A_{next} \Leftrightarrow Precs_a \in WS \ \& \ (\forall l \in L_p \mid next_l = id_a,$$
$$\exists \ ap \in A_{prev} \mid (id_{ap} = prev_l \ \& \ label_{ap} = DONE))$$

The $UPDATE\_PLAN$ operator updates the state of each action of the plan and the $GET\_NEXT\_ACTIONS$ operator returns the actions in A_{next}.

4.6.2 Action selection

Once the robot knows which actions need to be executed, it needs to choose one from the set of actions it can execute. To do so, it uses the $SELECT\_ACTION\_TODO$ operator which returns the action with the highest priority:

$$\underset{a \in \{A^R_{next} \cup A^X_{next}\}}{\text{argmax}} \quad priority(a)$$

Priorities used: In our case, we have chosen to give higher priority to the actions allocated to the robot compared to those not yet allocated. In general for this work, we have made the choice to postpone as much as possible the decisions to be made by the robot. Indeed, this is done in order to give as much latitude as possible to the human, which allows him to take the initiative until the last possible moment. In the current implementation of our system, the priorities of the different actions of the robot are the same, so the robot will simply select one. However, there is a possibility to later integrate costs as, for example, select the action the farthest of what the human is currently doing. Concerning the priorities of not allocated actions, we still follow the principle to postpone as much as possible the robot decision. To do so, we put a higher priority on what we call *analogous* actions. Two actions are considered *analogous* when they have exactly the same decomposition (same action name and same parameters). Indeed, as there is there is an action that has to be achieved several times, putting a higher priority on *analogous* actions allows to robot to execute one (and advance in the plan) while letting to the human the possibility to perform the other one.

4.6.3 Action allocation

Once a not yet allocated action is selected, the robot needs to decide if it should execute it or not ($ALLOCATE\_ACTION$ operator). To do so, the robot first looks for the possible actors of this action: agents which verify the preconditions of the action and which are not already busy. Note that even if the action was not allocated by HATP it is possible that there is only one possible actor. In this case, the robot automatically allocates the action to this agent. For example if the human is currently busy and there is a not allocated action to perform, the robot will execute it. If there is more than one possible actor for the action, the robot follows the algorithm 8.

Algorithm 8 Action allocation: $SP \leftarrow ALLOCATE\_ACTION(SP, a, WS, Prefs)$

Require: $a \in A^X_{next}$
  if $cost(a, R) << cost(a, H)$ **then**
    $actor \leftarrow robot$
  else if $cost(a, H) << cost(a, R)$ **then**
    $actor \leftarrow human$
  else if mode = negotiation **then**
    $answer \leftarrow ASK(a, H)$
    if answer = yes **then**
      $actor \leftarrow human$
    else
      $actor \leftarrow robot$
    end if
  else
    $actionPerformed \leftarrow WAIT\_ACTION(a, t)$ ▷ *adaptation mode*
    if $actionPerformed$ **then**
      $actor \leftarrow human$
    else
      $actor \leftarrow robot$
    end if
  end if
  return $actor$

First, the robot compares the estimated cost for the human and for itself to perform the action. If it considers it significantly more costly for the human to perform the action, it will allocate the action to itself. Then, we have developed two possible modes for the robot. In the first mode, called **negotiation** mode, the robot directly asks its human partner if he wants to perform the action and then allocates the action according to his answer. In the other mode, called **adaptation** mode, the robot waits a certain amount of time, and, if the human does not take the initiative to perform the action, it executes it.

Allocating an action to an agent can lead to other actions being automatically allocated. For this reason, after each allocation of an action, a new plan is built taking into account the possible allocations of the actions remaining in A_X.

Costs used for action selection ($cost(a, R)$ and $cost(a, H)$): In the current implementation, we use a cost concerning the *analogous* actions (see description in the previous subsection). These *analogous* actions will have a lower cost for the robot to execute them leading the robot to automatically execute one of them. Indeed, as there is several time the same action to perform, the robot can execute one while letting to the human the possibility to perform the other(s).

Other costs based on human preferences or on the context can be considered. For example, we can imagine having a list of actions the human likes to perform and another he dislikes. The robot can then allocate the actions following these preferences.

4.6.4 Action execution

Once the robot has decided to execute an action ($EXECUTE$ operator), it needs to be able to deal with the *similar* objects introduced before. To do so, we keep the principle that the robot waits until the last moment to take a decision. For example, if the robot and the human have to pick objects and place them in one out of several similar placements, the robot will first pick an object and only after choose a placement to place it. Then, when the robot has to choose an object, it will choose the one it considers the less costly.

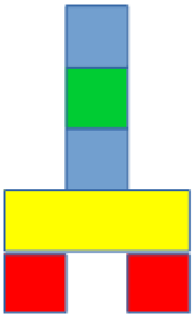
Finally, if the human approaches an object which is involved in the current robot action (e.g. if he places an object in a placement the robot has chosen), the robot first halts its action. Then, the robot looks if it can find another *similar* object. If it finds one, it continues its action with this object. If not, it waits for the human to retreat from the object, and if the human actions did not lead to a new plan it continues its action if possible.

Costs used for objects selection: Here we choose to put a lower cost on objects accessible only by the robot (we still want to let the maximum choices to the human). Then, we use a simple cost based on distance. In our cost, we get the distances between the agents hands (here we have chosen the right hand) and objects. For objects accessible only by the robot the costs will be proportional to the distance between the robot hand and the objects, leading the robot to choose the closest one. Concerning objects accessible also by the human, the costs will be inversely proportional to the distance between the human hand and the objects, leading the robot to choose the farthest object from the human to minimize the efforts for the human to reach the objects left.

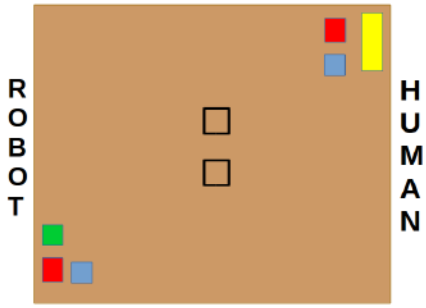
4.7 Results

4.7.1 Task

To illustrate the work done in this chapter, we use a task adapted to the manipulation abilities of a PR2 robot and inspired from the one in [Clodic 2014]. A human and a robot have to build a blocks construction as represented in Fig. D.11(a). At the beginning of the task, the robot and the human have several colored blocks they can access as in a set up like the one illustrated in Fig. D.11(b). Two identical placements are set on the table to indicate where to put the two red cubes.



(a) Goal of the task (side view)



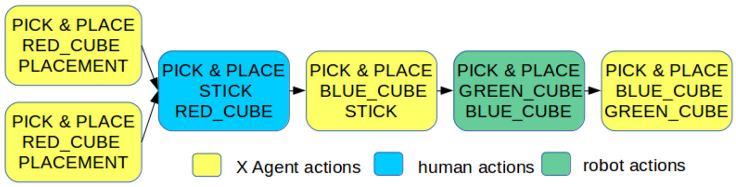
(b) One possible initial set-up (top view)

Figure 4.1: Description of the blocks building task. The human and the robot have to build the stack together. We assume that the robot and the human know where all the available blocks are. We would like the robot to adapt as much as possible to the human actions and decisions while avoiding useless or tiresome verbal interactions

4.7.2 Illustrative example

We will first present one possible scenario of the task described earlier which illustrates well the benefits of this work.

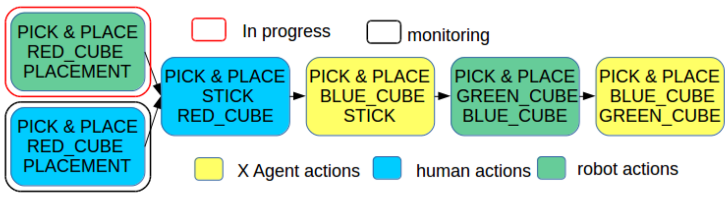
The presented scenario starts with the set-up in Fig. D.11(b). The plan produced by HATP for this set-up can be found in Fig. D.12(a). As this plan starts with two *analogous* actions for the *X agent* (place a red cube into a placement), the robot selects one and starts to execute it. So, the robot picks the red cube (Fig. D.12(b)) and, at the same time, looks for the consequences of its choice in the plan. As both agents own only one red cube, in the new plan computed by the robot (Fig. 4.2(d)) the human needs to place the second red cube. After picking its red cube the robot starts to place it on the placement to its right. However, the human picks his red cube and places it in the very same placement (Fig. D.12(c)). So, the robot stops its movement and adapts by placing its cube in the other placement (Fig. D.12(d)). Then, the human places the stick on the red cubes. In this scenario, we have chosen to set the robot into the **negotiation** mode. As the next action is allocated to the *X agent*, the robot asks the human if he wants to do it (*"Do you want to place the blue cube?"*). The human answers yes, leading the robot to compute the new plan in Fig. D.12(e) where the human have to place the first blue cube and the robot the second one. Finally, the human and the robot perform their last actions and achieve the goal.



(a) Initial plan

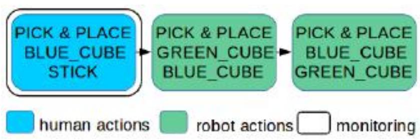
(b) The robot chooses to put the red cube in the placement to its right

(c) The human places his cube in the placement the robot chose



(d) Second computed plan

(e) The robot adapts by changing its placement choice



(f) Third computed plan

Figure 4.2: The human and the robot build a blocks construction together. The robot adapts its behavior to the human actions.

4.7.3 Quantitative results

In order to evaluate our system, we run it in simulation using the blocks building scenario. Different set-ups were used as initial state of the task: we randomized the number of cubes of each color in the environment and their position (accessible by the robot or the human). The robot was confronted to a simulated human with different types of behaviors. This simulated human performs all actions that are feasible only by him and answers robot questions. When confronted to an *X agent* action, he:

— chooses to perform it with 50% chance **(50%-case)**,
— systematically chooses to perform it **(hurry-case)**,
— systematically chooses not to perform it **(lazy-case)**.

Then, we settled two different human behaviors:

— **the "kind" human (case=K)** who adapts his behavior to what the robot verbalizes (ie does an action if the robot asks him and stops an action if the robot says it will perform it)
— **the "stubborn" human (case=S)** who does not react nor comply to robot verbalization (he will not change his decision whatever the robot says).

We compared 4 different modes:

— using the original system, called Reference System (RS), with all decisions and instantiations performed at planning time:
  — **RS- mode:** the robot verbalizes nothing (unless it is strictly necessary)
  — **RS-all mode:** the robot informs the human when he has to perform an action and when it will act,
— using the proposed system, called New System (NS):
  — **NS-N:** the robot uses the **Negotiation** mode previously defined when a decision need to be made concerning *X agent* action,
  — **NS-A:** the robot uses the **Adaptation** mode.

We measured:

— *the number of verbal interactions* between the human and the robot (either an information given by the robot or question asked), in Tab. 4.1.
— *the number of human/robot incompatible decisions*: either both decide to perform the same action (and the robot stops its own action to avoid the conflict) or both decide not to perform the action (the robot first asks the human to perform the action after a predefined time and, if after another period the human has still not executed the action, the robot looks for a new plan where it can proceed), in Tab. 4.1.

We also measured execution time but no significant difference was found between the different conditions. Indeed, this criterion is not pertinent here since, as all actions concern the same stack, they need to be performed one after the other. Consequently, there is no significant difference time between the different options.

Reference System performance: The verbalizations in the RS-none mode corresponds to the case where the human and the robot both choose not to execute the action: the robot tries to solve the conflict by asking the human to execute the action. Because the robot does not inform about its decisions, the number of verbal interactions is low in this mode. However, due to the same reason, there is several incompatible decisions in each conditions. The RS-all mode avoids incompatible decisions with the "kind" human. However, the number of verbal interaction is high

<table>
<tr><th></th><th>RS-none</th><th>RS-all</th><th>NS-N</th><th>NS-A</th></tr>
<tr><td>50%-K</td><td>0.6 (0.52)</td><td rowspan="3">0.0 (0.0)</td><td rowspan="6">0.0 (0.0)</td><td rowspan="6">0.0 (0.0)</td></tr>
<tr><td>hurry-K</td><td>0.3 (0.48)</td></tr>
<tr><td>lazy-K</td><td>0.9 (0.32)</td></tr>
<tr><td>50%-S</td><td>0.5 (0.53)</td><td>0.6 (0.52)</td></tr>
<tr><td>hurry-S</td><td>0.3 (0.48)</td><td>0.3 (0.48)</td></tr>
<tr><td>lazy-S</td><td>0.9 (0.32)</td><td>0.9 (0.32)</td></tr>
</table>

Table 4.1: **Number of incompatible decisions** between the human and the robot (i.e. either both agents decide to perform the same action or both decide not to perform a given action). Results for the reference system (RS) and the proposed system (NS-N for the negotiation mode and NS-A for the adaptation mode). The numbers correspond to means in 10 runs and their associated standard deviations.

<table>
<tr><th></th><th>RS-none</th><th>RS-all</th><th>Neg</th><th>Adapt</th></tr>
<tr><td>50%-K</td><td>0.4 (0.52)</td><td rowspan="3">6.0 (0.0)</td><td rowspan="6">1.2 (0.0)</td><td rowspan="6">0.0 (0.0)</td></tr>
<tr><td>hurry-K</td><td>0.0 (0.0)</td></tr>
<tr><td>lazy-K</td><td>0.9 (0.32)</td></tr>
<tr><td>50%-S</td><td>0.4 (0.52)</td><td>6.4 (0.52)</td></tr>
<tr><td>hurry-S</td><td>0.0 (0.0)</td><td>6.0 (0.0)</td></tr>
<tr><td>lazy-S</td><td>0.9 (0.32)</td><td>6.9 (0.32)</td></tr>
</table>

Table 4.2: **Number of verbal interactions**: question asked by the robot in the negotiation mode or an information given with the reference system. Results for the reference system (RS) and the proposed system (NS-N for the negotiation mode and NS-A for the adaptation mode). The numbers correspond to means in 10 runs and their associated standard deviations.

(6 as the number of actions to execute in the task and so to verbalize). With the "stubborn" human, even if the robot informs the human, incompatible decisions remains. The number of verbal interaction also increases as, when the human does not want to perform the action, as it is stubborn, the robot needs to compute a new plan where it executes the action, and so inform about the new action.

New System performance: We can see that the robot is able to avoid conflicts in all cases without being too talkative (or without being talkative at all for the adaptation mode). Moreover, the efficiency of the system is not degraded with the "stubborn" human: the system allows the human to execute the actions he wants without an increase of verbal interaction. Finally, here the adaptation mode performs better than the negotiation one since the human is simulated and always performs his actions in time. However, in a real context, the negotiation mode would certainly have the benefit to ensure the absence of conflicts even if the robot is a little more talkative. Moreover, a human would surely be more comfortable with a robot which directly asks when (and only when) there is a decision to take compared to a robot which has unnecessary waiting time. Such measure of "satisfaction" cannot be easily simulated and further experiments will be done with real humans.

4.8 Conclusion

In this chapter we have shown how we enable the robot to compute and execute more flexible Shared Plans. In these new plans, the needed decisions on who will execute an action and with which objects are let to the execution. A number of the presented algorithms involve cost estimation in order to decide between options. In the current system, simple costs are used but could be easily replaced by more elaborate ones. For instance, a finer estimation of action costs based on geometric reasoning and human efforts would allow better informed choice for action or object selection. Another interesting issue would be to integrate the estimation of accumulated costs of all actions remaining in the plan.

The benefits of this work have been demonstrated with an illustrative example and simulation results. We will show in Chapter 5 more complete simulation results which also include the work of the previous chapter as well as results with the system running in a real situation with real humans.

Chapter 5

Evaluation of the global system

Contents

5.1 Motivations

In the two previous chapters, we have presented several improvements on the way the robot elaborates and executes Shared Plans. We first endowed the robot with the ability to take into account humans mental state during Shared Plans execution. In a second time, we saw how the robot is able to compute more flexible Shared Plans where it identifies which decisions have to be taken at planning time and which one are better to be postponed. Then, the robot is able to take these decisions while smoothly adapting to the human choices.

These two improvements have been quantitatively and independently evaluated in simulation. In this chapter, we want to evaluate the global system including both extensions. Moreover, in addition to quantitative results, we want to evaluate the acceptance of the system by real users. To do so, we have defined a task which allows to highlight the benefits of the system. This task has been used to evaluate the global system in simulation in order to get quantitative results. Then, the same task (with minor modifications) has been used during a user study in the real robot in order to get a subjective evaluation of the global system.

The user study has been realized with the help of Camille Vrignaud, a master student in psychology.

5.2 Task

The task used for the global evaluation is inspired from the "Inventory scenario" of Chapter 3. In the task, the human and the robot have to scan several colored cubes and store them into a box of the same color. At the beginning of the interaction, both agents have a stack of colored cubes they can access (and only them can access). There are blue, green and red cubes. The stack of the human is located in another room, in a way that, to get an object, the human has to leave the sight of view of the robot (see Fig. D.13). For the cubes to be scanned, the agents need to put them on one of the two possible areas on the table in front of the robot (see Fig. D.13). Once a cube is placed on a scanning area, the robot can scan it by orienting its head and turning on a red light in the direction of the object (see Fig. D.15). If the robot scans an object while the human is not looking at him (e.g. he is in another room to pick a cube), the human will not be aware that the object has been scanned unless the robot tells him. Once the cube is scanned, it can be stored in a box of the same color (e.g. the blue cubes in a blue box). The robot has access to a blue box, the human to a green box, and both have access to a red box. Consequently, only the robot can store the blue cubes, only the human can store the green cubes and both can store the red cubes. As well as for his stack, the boxes of the human are located in another room (see Fig. D.13).

Both in simulation and in the user study, we compared 4 different conditions:

- using the original system, called Reference System (RS), with all decisions and instantiations performed at planning time and no estimation of the human mental state:
  - **RS-none mode:** the robot verbalizes nothing (unless it is strictly necessary)
  - **RS-all mode:** the robot informs the human when he has to perform an action, when it will act and about all actions he missed.
- using the proposed system, called New System (NS):
  - **NS-N:** the robot uses the **Negotiation** mode previously defined when a decision needs to be made concerning *X agent* action,
  - **NS-A:** the robot uses the **Adaptation** mode.

5.3 Evaluation in simulation

5.3.1 Modalities

We first evaluated our system in simulation. Different set-ups were used as initial state of the task: we randomized the composition of the stack of the human and the robot. In each case there was three cubes of each color (red, blue and green) and the robot stack was composed of 4 cubes and the human one of 5 cubes. The configuration of the cubes in the stacks were randomized. The robot was confronted to a simulated human with different kinds of behaviors. In all cases, the human was acting as below:

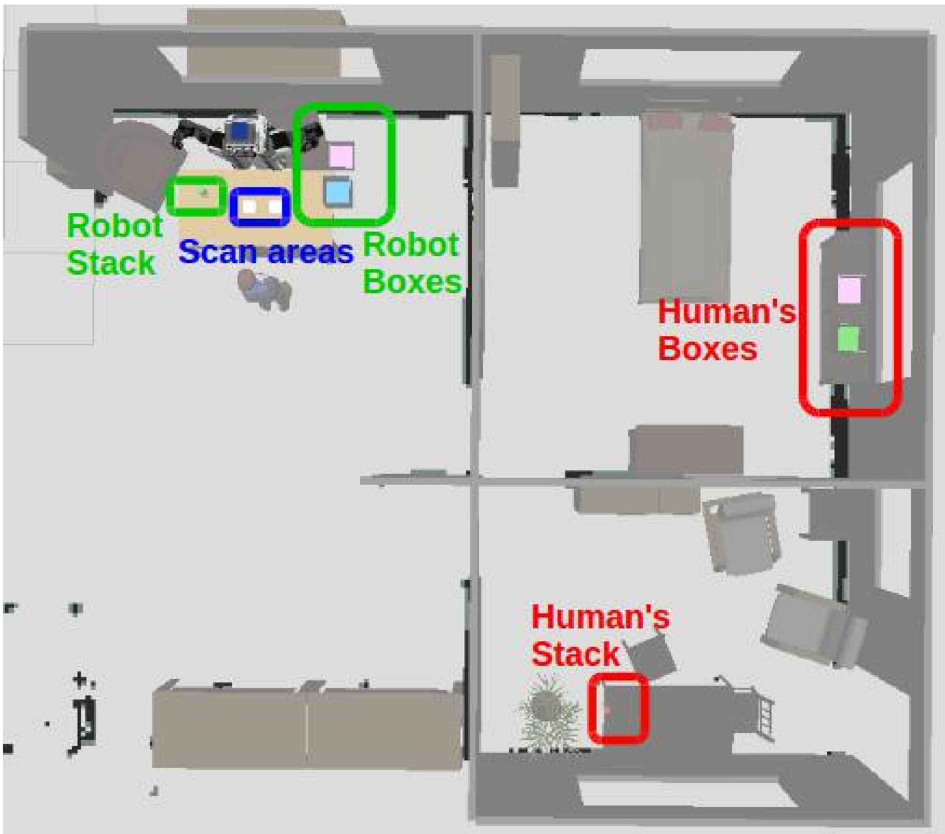


Figure 5.1: Set-up for the task used during evaluation. The human and the robot need to take the cubes from their stack and put them in the scan areas. Once a cube is in a scan area, the robot can scan it. Then, the agents can store the cubes in the boxes of the same color. The human has access to a green and a red box and the robot has access to a blue and a red box.

— when the human is in front of the robot with no cube in hand and there is a green cube he knows it is scanned, he goes to the boxes to store it.
— whenever the human is idle with no cube in hand, he goes to his stack to pick a cube and then comes back to the table.
— if the human has a cube in hand to be scanned, he put it on a scan area (if free). If there is no free area, the human waits in front of the robot.
— if the human has no more cube in his stack he waits in front of the robot.

When the human is in front of the robot with no cube in hand and there is a red cube he knows it has been scanned, the human chooses:

— to store it systematically (**hurry-case**)
— not to store it systematically (**lazy-case**)
— to store it with 50% chance (**50%-case**)

Then, we settled two different human behaviors:

— **the "kind" human (case=K)** who adapts his behavior to what the robot verbalizes (i.e. does an action if the robot asks him and does not execute

the actions the robot says it will perform)
— **the "stubborn" human (case=S)** who does not react nor comply to robot verbalization (he will not change his decision whatever the robot says).

In all cases, the human answers to the robot questions concerning the red cubes with the answer corresponding to his decision.

We measured:
— *the number of verbal interactions* between the human and the robot (either an information given by the robot or question asked), in Tab. D.2.
— *the number of human/robot incompatible decisions*: either both decide to perform the same action (and the robot stops its own action to avoid the conflict) or both decide not to perform the action (the robot first asks the human to perform the action after a predefined time and, if after another period the human has still not executed the action, the robot looks for a new plan where it can proceed), in Tab. D.1.
— *the total execution time:* for the human and the robot to perform the task, in Fig. D.14.

5.3.2 Results

	RS-none	**RS-all**	**Neg**	**Adapt**
50%-K	2.4 (0.84)	20.7 (1.34)	3.4 (1.51)	2 (1.33)
hurry-K	1.8 (0.79)	21.1 (2.08)	1.9 (1.10)	2.2 (1.13)
lazy-K	3.0 (1.33)	21 (1.56)	3.3 (1.42)	1.6 (1.17)
50%-S	2.5 (1.43)	23.9 (1.59)	3.3 (1.49)	1.7 (0.95)
hurry-S	1.5 (0.97)	20.9 (1.29)	2.4 (1.89)	1.9 (0.99)
lazy-S	3.2 (0.92)	25.2 (1.55)	2.8 (1.68)	1.8 (1.14)

Table 5.1: **Number of verbal interactions**: question asked by the robot in the negotiation mode or an information given with the reference system. Results for the reference system (RS) and the proposed system (NS-N for the negotiation mode and NS-A for the adaptation mode). The numbers correspond to means in 10 runs and their associated standard deviations.

RS-none performance: Even if the robot is supposed not to speak in this mode, we can see in Tab. D.1 that there are still verbalizations, especially in the cases where the human is lazy. These verbal interactions are due to two reasons. First, when the robot decides that the human should store a red cube and the human decides he will not do it, the robot unlocks the situation by asking the human to store the object (and in the stubborn case, as the human will still not do it, it then changes its plan). Secondly, when the human does not see that a cube has been scanned he will wait before storing it. As previously, the robot unlocks the situation by asking the human to store the object (as it detects that the human is not executing his action).

This mode is also the mode where there are the most incompatible decisions as the robot verbalize nothing. These incompatible decisions are mainly conflicts concerning the red cubes to store and the scan areas (as there is no notions of *similar* objects in this mode, the robot stops its actions if the human puts a cube in the same area it was aiming for even if the other is free).

Concerning the execution time, this mode is the one with the highest ones. The execution times are especially high in the stubborn and lazy cases, as, when the human decides not to store a cube, the robot wastes time to ask the human to do it and only then looks for a new plan where it stores the cube.

	RS-none	**RS-all**	**NS-N**	**NS-A**
50%-K	2.9 (0.99)	0.9 (0.57)	0.6 (0.7)	0.3 (0.48)
hurry-K	2.5 (0.97)	1.0 (0.94)	0.6 (0.52)	0.4 (0.52)
lazy-K	3.5 (1.08)	0.8 (0.63)	0.5 (0.7)	0.5 (0.53)
50%-S	2.9 (1.45)	1.9 (0.99)	0.6 (0.52)	0.5 (0.97)
hurry-S	2.3 (1.34)	1.0 (0.82)	0.5 (0.53)	0.4 (0.52)
lazy-S	3.5 (0.97)	2.6 (1.84)	0.3 (0.67)	0.4 (0.52)

Table 5.2: **Number of incompatible decisions** between the human and the robot (i.e. either both agents decide to perform the same action or both decide not to perform a given action). Results for the reference system (RS) and the proposed system (NS-N for the negotiation mode and NS-A for the adaptation mode). The numbers correspond to means in 10 runs and their associated standard deviations.

RS-all performance: In this mode, as expected, there is a lot of verbal interactions. Indeed, the robot informs not only about who should perform the actions but also about all actions that the human missed. However, even with the "kind" human, it is not enough to get rid of all conflicts. Indeed, there are still conflicts concerning the scan areas as the robot has to stop its action if the human puts his cube in the area it was aiming for. There are even more conflicts with the "stubborn" human as, even if the robot gives information, the human does not change his choices.

Concerning the execution times, they are low for the "kind" human as the human follows what the robot asks. However, with the "stubborn" human (not with the "hurry" one as the human takes the initiative to execute all possible actions), the execution times become higher. Indeed, when the robot has decided that the human has to perform an action, the robot wastes time to wait for the human to perform it before looking for another plan. These execution times are still lower than with RS-NONE because, as the robot informs for all missed actions, there is no time where the human waits to know that a cube has been scanned.

New system performance: We can see that the performance of the new system is globally better than in the two other modes. Concerning the incompatible decisions, it only remains the conflicts when the human puts a cube on the last available

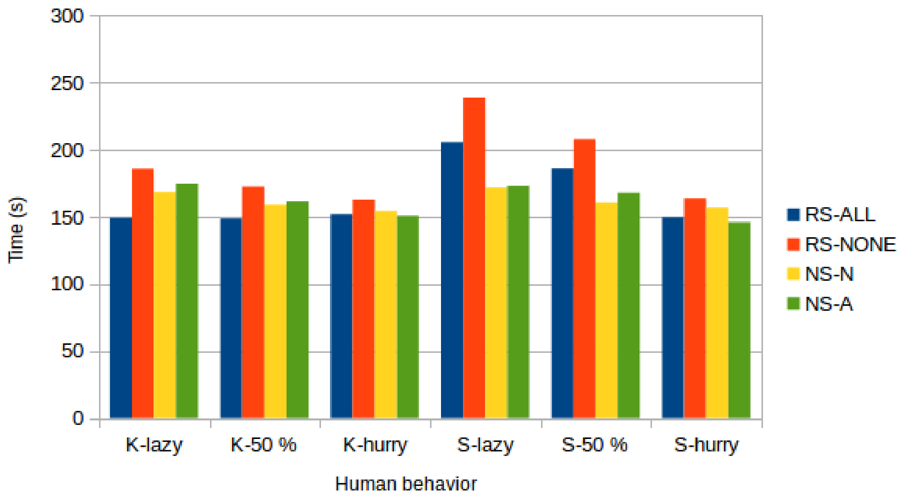


Figure 5.2: Time in seconds each system spent performing the task for each kind of human behavior (mean in 10 runs).

scan area and the robot was trying to put an object on it too. The execution times are lower than the reference system when the human is stubborn. Indeed, the robot does not wait for the human to perform actions he does not want to execute (it either asks or adapts). Moreover, as the robot informs the human about the cube which has been scanned during his absence (and which the human can store), the human does not wait to store cubes.

Concerning the verbal interactions, they are higher for the negotiation mode as the robot asks to the human if he wants to store the red cubes (but only when both agents are available). For the adaptation mode, these verbal interactions correspond to the information concerning the missing scan actions of the green or red objects.

5.4 User study

5.4.1 Background on evaluating human-robot interaction

The UX Model: The acceptability of complex technological systems as computer or robots is studied by researchers in social sciences. To do so they define what they call the User eXperience [Hassenzahl 2006] as:

> "a consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.)."

This definition has been designed as a model by [Mahlke 2008] as presented in Fig. 5.3.

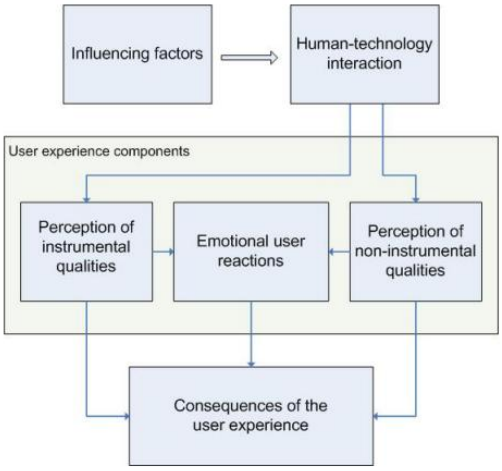


Figure 5.3: UX model by [Mahlke 2008]. The user experience is based on the human-technology interaction and is composed of three parts: the perception of instrumental qualities, the emotional user reaction and the perception of non-instrumental qualities.

This model is based on the human-technology interaction and is composed of three parts:

— **Perception of instrumental qualities:** The instrumental qualities of a technology, or also called pragmatic attributes, are strongly linked to the acceptability of the technology defined in [Dillon 2001] as:

> "the demonstrable willingness within a user group to employ information technology for the task it is designed to support."

One well known model of acceptability is the one presented in [Davis 1989] which can be found in Fig. 5.4. We can see on this model that the intention of the user to use the technology is linked to the perceived usefulness and ease of use (or usability) of the technology.

— **Perception of non-instrumental qualities:** In the definition of [Hassenzahl 2003], the non-instrumental qualities of a technology, or also called hedonic attributes, depend of the user and refer to the pleasure obtained by the use of the technology. It includes several notions such as the stimulation procured during the interaction, identification mechanisms and representations. This aspect is evaluated through the perception of the technology by the user. In the human-robot interaction context, the criteria to take into account are the esthetics of the robot, symbolic aspects and motivational aspects during the interaction.

— **Emotional user reaction:** The emotions of the user after the interaction with the technology will impact the final use of the technology. A positive emotion will support a future use while a negative emotion can lead to a reject of the technology.

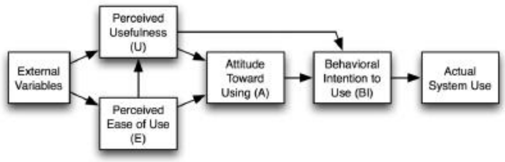


Figure 5.4: TAM model of [Davis 1989]. This model shows that the intention of use of the technology by the user depends of the perceived usefulness and ease of use of the technology.

Questionnaires to evaluate human-robot interaction: Several questionnaires have already been developed to evaluate human-robot interaction. [Hoffman 2013] allows to evaluate several aspects such as the "trust" in the robot or the "fluency" of the interaction. It has been used in several studies such as [Gombolay 2015] or [Dragan 2015]. However, this questionnaire does not deal with some aspects such as acceptability or usability. The Godspeed questionnaire [Bartneck 2009] allows to measure the perception of the robot by the human with questions relative to anthropomorphism or perceived intelligence. However, this questionnaire is focused on the evaluation of the perception of the robot and does not deal with the evaluation of the interaction and the usability of the system. The SUS (System Usability Scale) questionnaire [Brooke 1986] allows to measure the interaction of a user with an electronic system with 10 claims that subjects need to evaluate using Lickert scale from "totally agree" to "totally disagree". On the contrary of the Godspeed, the SUS questionnaire measures the usability of the system but lacks of measure concerning the perception of the robot or the interaction. [Heerink 2009] presents a toolkit to measure acceptance for assistive social robots. This toolkit is based on the UTAUT (Unified Theory of Acceptance and Use of Technology) questionnaire [Venkatesh 2003]. It has been well conceived in order to evaluate the perception and usability of the robot and more particularly for social robots. However, the questionnaire is more oriented toward the perception of the robot than the interaction and the collaboration. Finally, several studies as [Heerink 2010, Fischer 2016] use "homemade" questionnaire conceived for their experiment and not always easily reusable.

5.4.2 Construction of a new questionnaire

Why do we need it? We saw in the previous subsection that several questionnaires already exist to evaluate human-robot interaction. However, they mainly focus either on some specific basic behaviors or on evaluating human-robot interaction without concrete physical interaction. Even if these questionnaires are interesting in their respective fields of application, when it comes to evaluate high level decisions of a robot, few works have been done in the subject.

The decisions that we consider here correspond to physical (e.g. pick&place ..) and verbal actions that are involved when a human and a robot have to satisfy a joint goal: what action to perform, who will do and when. Indeed, when the task is a little complex, when various ways exist to achieve a same goal or when the spatial resource itself is shared by the human and the robot, it is important for the human to have a robot partner:

— that tracks and is permanently aware of the current state of the task
— that can comply with the human decisions
— that makes explicit its internal state
— that does all this with minimal intrusive behavior

The objective of researchers who contribute to the development of the robot high level decision abilities for Human-robot joint action is to come up with a robot that is able make the right decisions at the right time.

We need here a questionnaire which allows to evaluate the pertinence of high level decisions of the robot during human-robot Joint Action. We want the questionnaire to evaluate all aspects of the user experience in this context as well as specific aspects linked to the robot high level decisional abilities.

The questionnaire: In order to evaluate the user experience concerning the robot and the interaction, we have chosen to build a questionnaire where subjects have to place themselves in a self-assessment scale. This kind of questionnaire is often used in HRI because it allows to have quantitative measures on which it is possible to make statistical analysis. We organized the questionnaire on several dimensions, each one allowing to measure a specific aspect of the interaction. The first four dimensions allow to measure the different components of the UX model presented before:

— *Evaluation of perception of the instrumental qualities:* we constructed two different dimensions of the questionnaire based on the TAM model (Fig. 5.4). The first one, based on [Weistroffer 2014] and called the **Collaboration** dimension afterward, allows to evaluate how the subject perceived the utility and the usability of the interaction (ergonomic criteria). The second one, based on the French version of the AttrakDiff questionnaire [Lallemand 2015] and called the **Interaction** dimension afterward, allows to evaluate the behavioral intention of use.
— *Evaluation of the perception of non-instrumental qualities:* we based this part on the Godspeed questionnaire [Bartneck 2009]. It allows to evaluate how the human perceived the robot in general. The associated dimension of the questionnaire will be called the **Robot perception** dimension afterward.
— *Evaluation of the emotional user reaction:* we used the AffectButton from [Broekens 2013]. The associated dimension of the questionnaire will be called the **Emotions** dimension afterward.

In addition to these dimensions, we have added two other dimensions that are more specific to the context of high level decision and physical human-robot Joint Action:

— **Verbal:** this dimension allows to evaluate how the human perceived the verbal interaction with the robot (did the robot verbalized the good information at the right time).
— **Acting:** this dimension allows to evaluate how the human perceived the decisions of the robot concerning its actions (did the robot choose to perform the right actions at the right time).

Concerning the emotions dimension of the questionnaire, as said previously, we use the AffectButton from [Broekens 2013]. Subjects have to choose between five emoticons the one which corresponds the most to their feelings. Concerning the other dimensions, several antonym items are used by dimension (between 3 and 8). Subjects have to answer a question by placing themselves in a scale of 100 between these antonym items. The English translation of the questionnaire can be found in Appendix C.

5.4.3 Adaptations of the task for the study

Before realizing the real study, we have made some pre-tests by running the task in the robot with few subjects. During these pre-tests we noticed several possible problems that we fixed by proceeding to small adaptations of the task.

Introduction of a red video tape box: In certain cases, the configuration coupled to the decision of the subjects led to not having any decision in the task concerning the red cubes. Indeed, there were cases where, each time there was a red cube to store, one of the two agents were busy (either the human was in another room to pick or store an object or the robot was performing another action).

To ensure that, at each interaction, there is at least one decision to take between the human and the robot, we added to the objects to scan and store a red video tape box. The human and the robot both have a red video tape box in the same placement as their stacks of cubes. At the end of the task, when all the cubes are scanned and stored (and so both agent are available), **only one** of the two video tape boxes (the one of the human or the one of the robot) needs to be put on a scan area. Then, as well as for the cubes, the robot scans the video tape box. Finally, as the video tape box is red, it needs to be stored in a red box either by the human or the robot.

Distraction task: We noticed that some subjects tried not to miss any action of the robot (they stayed in front of the robot each time there was a cube to scan and they hurried in the places where they cannot see the robot). Consequently, there was not missing knowledge during the task for these subjects. To ensure that all subjects miss some actions of the robot, at one predefined point of the task, the experimenter asked the subject to leave the task for a while to perform another

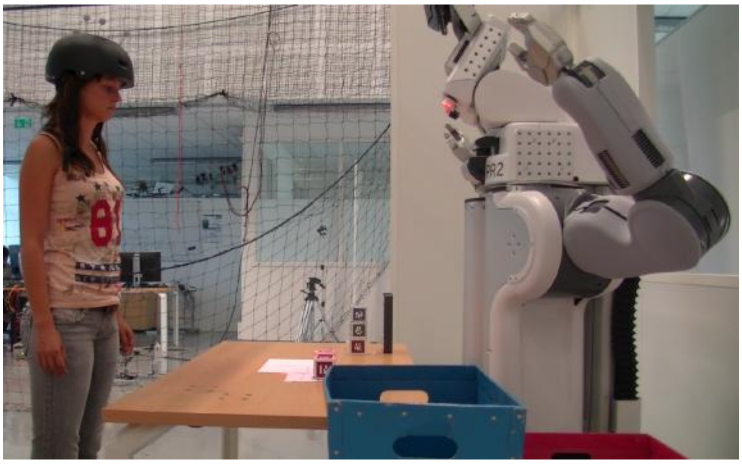


Figure 5.5: The PR2 robot interacting with a subject to achieve the task. The robot is scanning the cube before storing it.

task. In this task, the subject has to build a construction shown in a picture with Lego bricks. Once the construction achieved, the subject is free to go back to the main task.

5.4.4 Questionnaire and protocol

Each subject of the study had to interact with the robot to achieve the task previously described, and in the four conditions described in Sec. 5.2. The order in which they were confronted to the different conditions was randomized. There were four different configurations for the stacks of the human and the robot. The attribution of each configuration to a condition was also randomized for each participant.

At their arrival, the participants were introduced to the robot and the environment of the study by the experimenter. Then, participants were asked to read instructions explaining the task and its constraints. The experimenter checked the good understanding of the instructions and showed the placements of the different objects of the task. The participants were then asked to perform a quick familiarization task. In this task, the human and the robot had only one cube in their stacks (a blue for the human and a green for the robot). They had to put them in the scan areas, scan them and then store them in the appropriate boxes. There was no video tape box in the familiarization task.

After each interaction with the robot (for each condition), the participants were asked to fill the questionnaire presented above. In addition to this questionnaire, after each interaction with the robot (including the familiarization task), we asked participants to answer a small yes/no questionnaire. This questionnaire contains general questions about what happened during the interaction (e.g. "Do you think all the cubes have been scanned?"). The aim of this questionnaire was to remind the key points of the interaction to the subjects (because we noticed during the pre-tests that subjects were kind of "lost" and did not know on what to focus their attention). This questionnaire can also be found in Appendix C.

5.4.5 Hypothesis

The results expected of the user study are:

— **Hypothesis 1:** The new system will be preferred to the old one by the users.

— **Hypothesis 2:** For the new system, the negotiation mode will be preferred by the user to the adaptation one. Indeed, even if in simulation better results where found for the adaptation mode in Chapter 4, we strongly believe that naive users will be more comfortable with a robot asking whenever there is a choice.

5.4.6 Results

21 subjects took part in the study (8 women and 13 men). They were all fluent in french and had no significant experience in robotics.

Questionnaire validation: In order to validate the coherence and uniformity of the questionnaire used during the study, we calculated Cronbach's alpha for each dimension of the questionnaire [Cronbach 1951]. We calculated these values for the RS-none condition which is the closest of based condition. These values can be found in Tab. 5.3. To consider that the coherence of a dimension is validated, alpha should be of 0.7 or higher. We can see that all dimensions of questionnaire (the french version) are validated here.

Dimensions	Cronbach's alpha
Verbal	0.73
Acting	0.85
Collaboration	0.76
Interaction	0.9
Robot perception	0.84

Table 5.3: Cronbach's alpha for the different dimensions of the questionnaire. An alpha of 0.7 and higher means the dimension is validated.

Concerning the scores of the questionnaire, the total results for the questionnaire evaluating the subjects feeling concerning the robot and the interaction can be found in Fig D.16 and the details for each dimension in Fig 5.7. We will discuss the results here below.

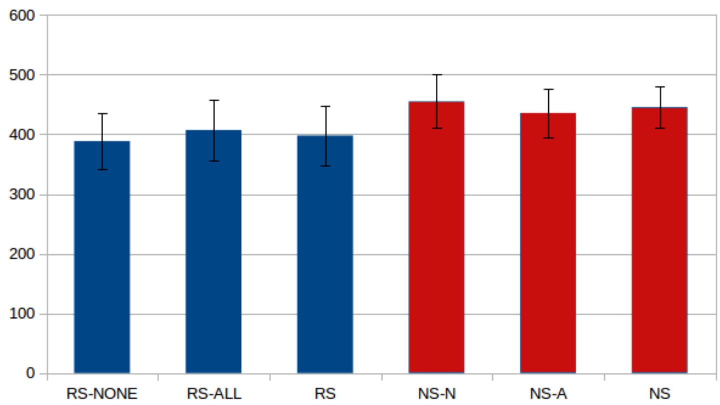


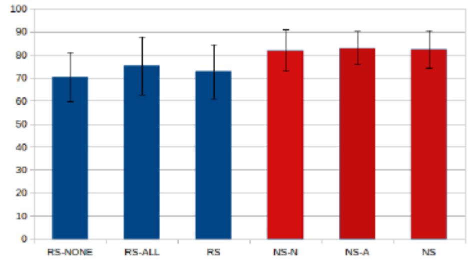
Figure 5.6: Total scores of the questionnaire. It is calculated by adding the scores of all dimensions previously harmonized in a 100 scale. The different modes are for the reference system RS-all when the robot verbalizes everything and RS-none when the robot verbalize nothing and for the proposed system NS-N for the negotiation mode and NS-A for the adaptation mode. The RS and the NS columns are the means respectively for the Reference System and the New System.

Comparison of the two systems: We first compared the Reference System (RS) to the New System (NS) in order to validate the first hypothesis. We compared the two systems looking at the total score of the questionnaire and each dimension individually. We applied student T-test when the data where normally distributed and Wilcoxon tests when the data where not normally distributed. The obtained results can be found in the first column of Tab. 5.4.6. We can see that for the total of the questionnaire and for all dimensions except the Verbal one, the new system has been evaluated significantly better than the reference system ($p < 0.05$). Consequently, we can consider the first hypothesis as validated. The difference was particularly visible for the *Acting* dimension of the questionnaire ($p \simeq 0.003$). It shows that the algorithms developed for the robot to be able to take the appropriate decisions at the right time during Shared Plan achievement have been appreciated by the subjects. The reasons why no significant difference was found for the verbal dimension will be explained below.

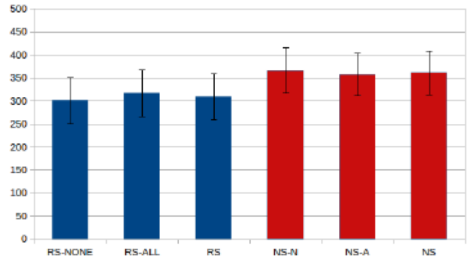
Dimension	**RS/NS**	**NEGO/ADAPT**
Total	(W) p = 1.002e-3*	(W) p = 0.147
Emotion	(W) p = 2.513e-2*	(W) p = 0.830
Collaboration	(T) p = 4.321e-3*	(W) p = 0.434
Interaction	(T) p = 3.605e-2*	(W) p = 0.237
Robot perception	(W) p = 2.080e-2*	(W) p = 0.531
Verbal	(T) p = 0.308	(W) p = 8.966e-3*
Acting	(T) p = 3.537e-3*	(W) p = 0.222

Table 5.4: P-values from the student T-tests and Wilcoxon tests. The first column corresponds to the comparison of the Reference System (RS) and the New System (NS). The second column corresponds to the comparison of the negotiation and the adaptation modes of the new system. (T) means a student T-test has been applied and (W) means a Wilcoxon test has been applied. * means the difference between the results is significant (p-value < 0.05).

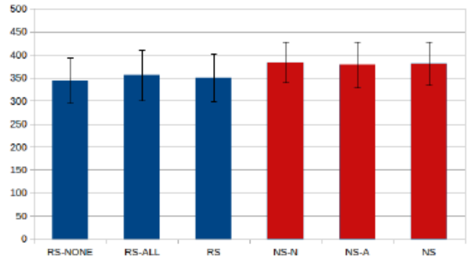
Comparison of the negotiation and adaptation modes: We then compared the negotiation and the adaptation modes of the new system in order to validate the second hypothesis. We compared the two modes looking at the total score of the questionnaire and each dimension individually. We applied student T-test when the data where normally distributed and Wilcoxon tests when the data where not normally distributed. The obtained results can be found in the second column of Tab. 5.4.6. Here we can see that, even if the means of the negotiation mode are higher than the ones of the adaptation mode, no significant difference was found except for the Verbal dimension. Indeed, the principal difference between the two modes consists in the way the robot deals with the choices concerning the red objects. In the negotiation mode the robot asks to the human if he wants to perform the action and in the adaptation mode the robot does not speak and simply adapts. We can deduce that maybe the naive users of the study preferred the negotiation



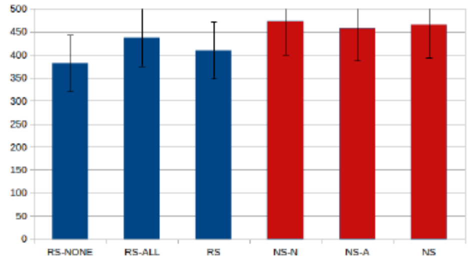
(a) Scores for the "Emotion" dimension of the questionnaire. The scores have been put in a 100 scale for harmonization (each emoticons corresponding to a score of 20, 40, 60, 80 or 100).



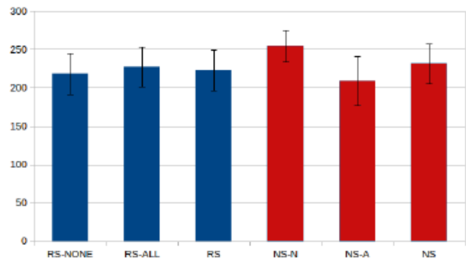
(b) Scores for the "Collaboration" dimension of the questionnaire



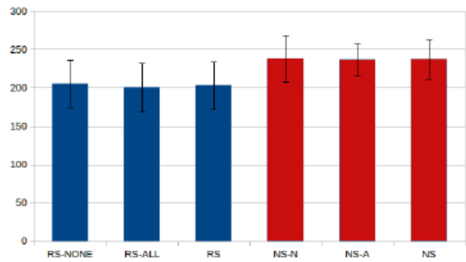
(c) Scores for the "Interaction" dimension of the questionnaire



(d) Scores for the "Robot perception" dimension of the questionnaire



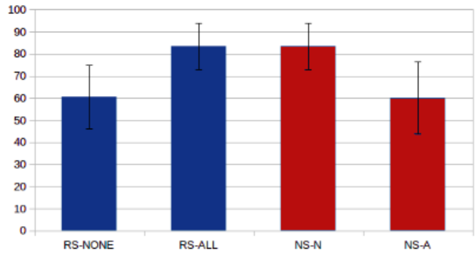
(e) Scores for the "Verbal" dimension of the questionnaire



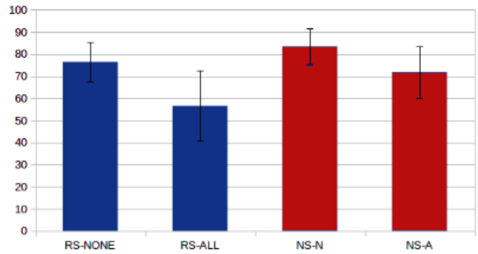
(f) Scores for the "Acting" dimension of the questionnaire

Figure 5.7: Results on the questionnaire evaluating the subjects feeling concerning the robot and the interaction given to the participants during the user study. The different modes are for the reference system RS-all when the robot verbalizes everything and RS-none when the robot verbalize nothing and for the proposed system NS-N for the negotiation mode and NS-A for the adaptation mode. The RS and the NS columns are the means respectively for the Reference System and the New System.

mode mainly because it was comforting to have the robot asking when there was a choice. In conclusion, looking at these results and previous results in simulation, maybe the negotiation mode should be preferred for first or punctual interactions with the robot, and, when the user becomes more used to the robot, the adaptation



(a) Score for the question concerning the verbal interactions where subjects were asked to choose between "insufficient" (0) and "sufficient" (100)



(b) Score for the question concerning the verbal interactions where subjects were asked to choose between "superfluous" (0) and "pertinent" (100)

Figure 5.8: Details of the results on the verbal dimension of the questionnaire. The different modes are for the reference system RS-ALL when the robot verbalizes everything and RS-NONE when the robot verbalize nothing and for the new system the NEGO mode and the ADAPT mode.

mode may be preferred.

Focus on the verbal dimension: We applied a Friedman ANOVA test to compare the different conditions of the Verbal dimension. The negotiation mode of the new system had scored significantly higher than the RS-NONE condition and the adaptation mode ($p < 0.05$). Even if the negotiation mode had a higher score than the RS-ALL condition, the difference was not found significant. Indeed, when discussing with subjects after the experiment, we found that, for some of them, as they felt stressed, the fact that the robot was speaking a lot was comforting because they did not have to take decisions nor to interpret robot actions. However, they also point out the fact that, even if they found it comforting the first time, if they had to interact with the robot several time in this mode, they would quickly find it "annoying". Indeed, if we look at some details of the questions asked into the verbal part of the questionnaire, the verbal interaction of the robot have been found more superfluous in the RS-ALL mode than in the other modes (see Fig. 5.8(b)). Moreover, the verbal interactions in the RS-NONE mode and the adaptation mode have been found less sufficient than in the other modes (see Fig. 5.8(a)). Indeed, the fact that the robot does not inform about its choices (and more particularly concerning the red objects) was found disturbing by the participants.

5.5 Conclusion

The aim of this chapter was to evaluate the algorithms presented in the last two chapters in order to improve the Shared Plan elaboration and execution by the robot. The new system, with its two possible modes (negotiation and adaptation) has been compared to a reference system corresponding to the state of the art before the ameliorations (with two possible options for verbalization). The evaluation has been done both in simulation and with a real study in the real robot.

Both evaluations have shown that the new system performs better than the old one. In simulation, the adaptation mode performed a little better than the negotiation mode (a little less verbalizations). However, the naive users during the user study preferred the negotiation mode mainly because it was comforting to have the robot asking when there was a choice. In conclusion, maybe the negotiation mode should be preferred for first or punctual interactions with the robot, and, when the user becomes more used to the robot, the adaptation mode should be preferred.

Moreover, during the user study, we constructed a questionnaire in order to evaluate users feeling concerning the collaboration with the robot which has been validated (in term of intern coherence) thanks to the study data. This tool is generic enough to be used for other studies where a robot collaborates with a human.

The user study presented in this chapter also allows us to get more insights on experiments with naive subjects. Indeed, there is a real difficulty to correctly evaluate works in decision for human-robot interaction due to several reasons:

— one of the first challenge is to find a task for the human and the robot to perform together. Indeed, this task should be sufficiently interesting to allow the evaluation of the system but not too complex if we do not want the subject to focus too much on the task rather than on the robot behavior. Moreover, the task should be adapted to the perception and manipulation abilities of the robot.

— another difficulty when the robot interacts with a human is to isolate the decisional aspect from the other robot abilities. Indeed, we figured out that, because they are not used to interact with robots, it is very hard for the subjects to distinguish two different behaviors of the robot (even if the difference seems huge to roboticists) because they have too much things to observe in the robot behavior.

— as the implementation of the robot decisional abilities usually relies on the other components of the robotics architecture (e.g. perception, manipulation), it is complicated and time consuming to obtain a global system sufficiently robust for a user study.

— finally, as the subjects during the user studies are naive, we can question the obtained results. Indeed, we have experimented the fact that it is difficult, in a one session, while trying to avoid introducing heavy biases, to ask naive users to distinguish between a one-shot use of a robot and its potential daily use. Then "annoying", "repetitive", "intrusive", "delayed" , "lacking of fluidity", "superfluous" behaviors will be certainly more severely evaluated. For the time being, the users are basically happy to "play" with the robot. This was perhaps enforced by the fact that in all the versions that we have proposed to the users, since the robot observes correctly the state and produces valid plans, the tasks is always finally achieved. Some UX models take into account the temporal aspect of the interaction[Kim 2015], however, it usually implies to make long term user studies which are not easily feasible in our research context. One challenge would be to find a way to evaluate this aspect without a need of a long term study, maybe by performing user studies with subjects who are not naive but more used to robotics systems.

Part III

Other contributions to Human-Robot Joint Action

Chapter 6

Non-verbal communication: what should the robot do with its head?

Contents

6.1 Motivations

During human-robot Joint Action, the robot needs to provide to its human partner a lot of information on what it is doing, understanding and what should be done next. All this information needs to be communicated without being too chatty by verbalizing everything. To do so, humans usually rely on non-verbal communication [Ekman 1969, DePaulo 1992]. In order to collaborate in a fluent and natural way with humans, the robot needs to exhibit a non-verbal behavior adapted to the Joint Action. Non-verbal behavior can come from multiple sources (e.g. facial expressions, posture, head behavior). In this chapter we want to focus on the robot head behavior. Head behavior means here the orientation of the robot head which reflects what the robot is looking at [Imai 2002]. In a first step, we used the state of the art and bibliography concerning humans behavior in order to identify which kind of behaviors or signals are needed during human-robot Joint Action. We studied in more details some of these signals with the help of an online questionnaire. Then, we investigated how to combine all these signals and behaviors in order to produce a robot head behavior which supports Joint Action. This work has been done in collaboration with another PhD student Yoan Sallami.

6.2 Background

6.2.1 On the use of gaze

Non-verbal communication between humans has been deeply studied in social sciences literature [Ekman 1969, DePaulo 1992]. Studies have shown that, when we are in the context of a social interaction, we adapt our non-verbal behavior in order to better coordinate with our partner(s) [Becchio 2010, Vesper 2010]. Non-verbal behavior can come from several sources. Posture can be used to communicate a lot of information [Mehrabian 1969] as well as facial expressions [LaBarre 1947]. Gaze is also very important in non-verbal communication. During social interaction, people look at others for an average of 61% of the time [Argyle 1972]. Several kinds of gaze behavior can be identified [Mutlu 2009a]:

— **One-sided gaze, looking at:** A looks at B in or between the eyes, or, more generally, in the upper half of the face [Cook 1977].
— **Mutual gaze, eye contact:** Both A and B look into each other's face, or eye region, thus acting simultaneously as sender and recipient [Von Cranach 1973].
— **Gaze avoidance:** A avoids looking at B especially if being looked at, and/or moves the gaze away from B [Von Cranach 1973, Emery 2000].
— **Gaze following:** A detects B's direction of gaze and follows the line of sight of B to a point in space [Emery 2000].
— **Joint attention:** A follows B's direction of attention to look at a fixed point in space (such as an object) [Butterworth 1991].
— **Shared attention:** Both A and B look at a fixed point in space and are aware of each other's direction of attention [Baron-Cohen 1995, Emery 2000].

These behaviors allow to support the interaction in several ways:

— **Support dialogue and turn-taking:** Research on conversational functions of gaze show that gaze behavior is closely linked with speech [Argyle 1976]. Gaze is more specifically used to indicate the beginning and the end of an utterance in order to facilitate the switch of roles [Kendon 1967] and the organization of a group discussion [Goffman 1979].
— **Support action understanding:** Studies have shown that, an individual acting in a social context will not act in the same way than when he is acting alone [Becchio 2010, Vesper 2010]. More particularly, the use of gaze allows to better infer people's motor intention [Castiello 2003, Pierno 2006].

— **Support mental states:** Social and developmental psychological studies have shown that through observing others' gaze patterns, people infer personality traits [Kleinke 1986] and detect and infer deception [Hemsley 1978]. Moreover, gaze can also be used to support perspective taking by analysing the object an individual is looking at [Furlanetto 2013] and, in certain cases, can be more efficient than language to pass information [Neider 2010].

6.2.2 On the use of the head in robotics

Several works studied the use of non-verbal behaviors in robotics. In [Breazeal 2005], Breazeal et al. have shown that people infer task-relevant mental states of a robot not only from explicit social cues that are specifically intended to communicate information to the human (e.g., nods of the head, deictic gestures, etc), but also from implicit behavior (e.g., how the robot moves its eyes: where it looks and when it makes eye contact with the human). Studies have also shown that explicit non-verbal behavior such as pointing an object we are referring to helps for a better understanding [Häring 2012, Salem 2011].

Concerning the gaze, [Imai 2002] established that, in absence of eyes independent from the head, the perception of the robot gaze is coupled to the robot head orientation. Many works studied the robot head behavior during conversation. They showed that looking at the human at the right time helps to take turn [Boucher 2010, Skantze 2014] and that looking at an object the robot is referring to helps the understanding of the human [Mutlu 2009b, Staudte 2011].

Concerning the use of the head during Joint Action, [Lallée 2013] have shown that, in the absence of language, the use of a head behavior based on a known Shared Plan helps the coordination between the human and the robot. [Zaga 2017] have shown that, with a robot producing "social-gaze movements" (tracking co-player actions), perception of animacy and likeability significantly increases. Moreover, if we add a "deictic gaze" (providing helpful referential information for the completion of the task), perception of helpfulness significantly increases. The importance of the head behavior during human-robot Joint Action has also been studied in [Boucher 2012] where they have illustrated that humans are able to make anticipatory decisions based on robot gaze cues.

Several works studied what the robot should look at during handover [Moon 2014, Gharbi 2015]. They have shown that the appropriate gaze cues improve synchronisation and the perception of the efficiency of the handover. The robot should look at the place where the object will be exchanged in advance in order for the human to anticipate. Moreover, the handover is perceived as more natural if the robot looks at the human at the end. The use of gaze has also been studied during navigation tasks as approaching a person [Fischer 2016] or navigating between humans [Khambhaita 2016]. These works showed the importance of switching between the path to go and the human.

6.3 Brainstorming concerning the needed behaviors and signals

Previous works have shown that non-verbal behavior, and particularly gaze - head - behavior is key in human-robot interaction. Based on the works presented earlier and on our observations of human-human and human-robot interactions, we listed what we think are needed components of a robot head behavior appropriate to human-robot Joint Action. We organized these components in four "families" based on the robot activity.

The robot acts: Head behavior has a big influence on the legibility and the predictability of the robot actions. Previous works in robotics demonstrated its influence during handover [Moon 2014, Gharbi 2015] and studies on humans showed that actors should adapt their non-verbal behavior to the context of Joint Action [Becchio 2010, Vesper 2010]. Moreover, for some actions, the robot needs its head (where there are usually cameras) to perform the action (e.g. precise grasp of an object). Consequently, we strongly believe that having a head behavior consistent with the robot action not only in a functional point of view but also with the purpose of showing what the robot is currently doing is one first key component of an appropriate robot head behavior.

Moreover, we asked ourselves if the anticipation of the next robot actions with its head (for example by pointing the next target) can help the human to better predict and understand the robot next actions. This question will be addressed in the next section.

The robot speaks: Head behavior during dialogue has been deeply studied in psychology and several studies have been done in the subject in robotics. These studies highlight the importance of looking at the receiver when we speak especially at the end of an utterance [Boucher 2010, Skantze 2014]. The importance of looking at the objects we are referring to has also been demonstrated [Mutlu 2009b, Staudte 2011]. An appropriate robot head behavior should take into account these two aspects.

The robot observes: During Joint Action, the robot not only needs to be understood but also to show that it is attentive and able to recognize and track what its partner is doing. The robot head behavior should allow to show the interest of the robot in the human activity. Different ways to do it by tracking the human hand will be studied in the next section.

The robot should also be able to engage in a joint or a shared attention with a human. Indeed, if the human stares at the robot, the robot should return his look. In a same way, if the human stares at an object, the robot should be able to show its interest in the object too.

The robot coordinates: One important aspect of Joint Action is coordination. Previous works have shown that taking into account the Shared Plan when communicating helps the interaction [Lallée 2013]. To support the communication of the Shared Plan, the robot should provide the appropriate signals at the right time. These signals should facilitate turn-taking and promote the fluid execution of the Shared Plan. They will be deeply studied further in the next section.

6.4 Deeper study of some signals

We identified in the previous section some essential components of a robot head behavior adapted to human-robot Joint Action. In order to get some cues concerning some specific parts of this behavior, we performed an online video based study. This method has been tested in [Woods 2006] and its potential as a technique for prototyping, testing and developing HRI scenarios has been proved.

In the performed study, for each specific behavior or signal which we tested, we asked participants to watch several short videos of a human-robot interaction where the behavior/signal was declined in different forms. We asked subjects to compare the different videos by answering several questions. A total of 59 (30 women and 29 men) answered the questionnaire. All participants had no experience in robotics. The English version of the form (the form was available in French and in English) is available at `https://goo.gl/forms/q5tUcPHRJdqg8EMH3`.

The same task was used in each video and was explained at the beginning of the questionnaire to the participants. In this task, a human and a robot have to build a stack of colored cubes together. The cubes have to be stacked in a precise order (see Fig. D.17(a)) and, at the beginning of the interaction, several cubes are reachable by the human and the others by the robot (see Fig. D.17(b)).

In addition to the tested behaviors and signals, the robot had a "basic behavior":

— By default (when it has no specific target), the robot looks at the human head.

— When acting, the robot looks at the "target" of its action (i.e. the object to pick and the support where to place).

This behavior has been constructed based on bibliography and observations during previous experiments.

(a) Beginning of the interaction session.The blue and the green cubes are accessible by the human while the black and the red cubes are accessible by the robot.

(b) End of the interaction. The stack should follow a precise order (red, black, blue, green).

Figure 6.1: Task used in the on-line video based study. In this task a human and a robot have to build a stack of colored cubes. Different head behaviors of the robot was compared in the videos.

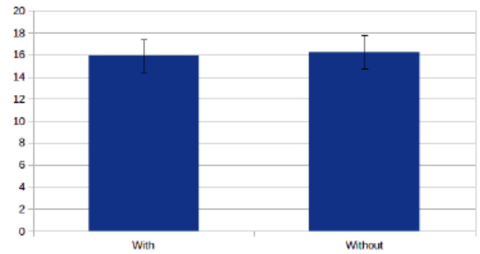
We grouped the different tested behaviors and signals into four sections:

— **Anticipation of robot actions:** we first studied whether or not the robot should anticipate its next action with its head.
— **Tracking human's activity:** we then studied in two stages how the robot should track the human's activity.
— **Helping the human to perform his next action:** we studied two kinds of signal aiming to help the human to perform his next action.
— **Finding the priority target:** finally, we studied a situation where the robot has two different targets to look at the same time and studied which one to prioritize.

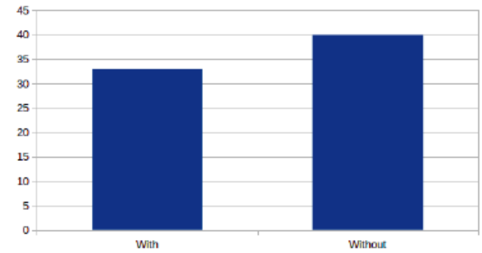
6.4.1 Anticipation of robot actions

The first behavior we studied concerns when the robot is performing actions. We asked ourselves if the fact that the robot anticipates its next action with its head would benefit to the Joint Action and more particularly to the understanding of the robot actions by the human. To do so, we asked people to watch two videos: one with anticipation and one without. These two videos are focused on the first part of the interaction session where the robot put the first two cubes on the stack. In the video without anticipation, the robot simply looks at the cube it is picking and the support where it is placing the cubes (as described in the "basic behavior"). In the video with anticipation the robot still looks at the cube it is picking and the supports where it is placing the cubes. However the robot starts looking at the second cube to put on the stack while it is retracting from its first action (putting the first cube).

After having watched the two videos, participants were asked to answer several questions for each video. There was 4 5-scales questions for each video concerning the predictability of the robot behavior and the fact that the robot head behavior



(a) Scores on the questions asked for each video of the scenario concerning the anticipation of robot actions. The score is the addition of the rates (in a scale of 5) of the 4 questions.



(b) Preferences for the scenario concerning the anticipation of robot actions. Numbers represent the number of times where a video was chosen for the preference question (on 59 participants).

Figure 6.2: Results for the scenario concerning the anticipation of robot actions. No significant differences was found between the two conditions (with anticipation and without anticipation).

was adequate, clear and useful to the interaction. Then, the participants were asked to choose which video they preferred between the two videos (with the possibility to select both). These questions can be found in Appendix B. We also let a free space for comments at the end of the questionnaire.

The results of the questionnaire for this scenario can be found in Fig. 6.2. No significant difference was found between the conditions either concerning the scores in the 5-scales questions ($p=0.3467$) or the preferences ($p=0.3711$). Indeed, based on the free comments of the participants, we found that they had difficulties to find the differences between both conditions, and, when they found the differences, they were sometime disturbed by the fact that the robot looks at an object without acting (in the condition with anticipation).

We can conclude that the way we implemented anticipation is not conclusive for Joint Action, or at least for the kind of task tested here. Indeed, one factor here is that the task and needed actions are well known of both participants (as the stack order is predetermined, there is only one action to do at each time). Consequently, the anticipation may not be needed here but may be more interesting in a scenario where the human does not know the next robot action. Moreover, maybe the way the anticipation behavior is implemented is not the best way to do it. Further investigations may be needed on this subject.

6.4.2 Tracking human's activity

We then studied how the robot should show that it follows and understands human's actions. We want here to find a behavior adapted to the information the robot has concerning the human. In our system, the robot detects the human with a motion capture system (see description in Introduction) and has information about the head and the right hand positions and orientations. Here we tried to find a relatively simple behavior which, with a minimum information allows the robot to show its attention and understanding.

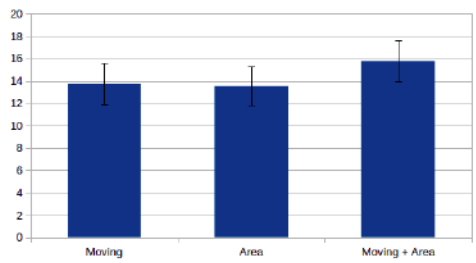
What to look: First, we compared different ways to decide when to look at the human's hand or when to look at the human's head. We compared three different conditions with videos based on the end of interaction session (the part where the human puts the two last cubes). In all videos the human puts the two cubes one after the other. He also records a stop time during his first action by hesitating a short moment. Then, the different conditions were:

— **Moving:**the robot looks at the human's hand whenever the hand is moving and looks at the human's head when the hand is not moving (with a small hysteresis in order to avoid going back and forth between the hand and the head).

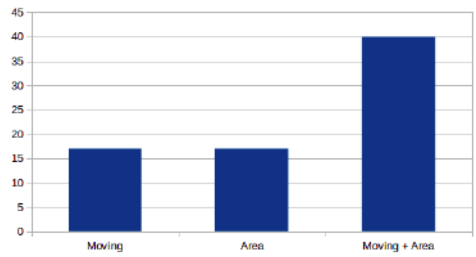
— **Area:** the robot looks at the human's hand whenever the hand is into a "working area" and at the human's head whenever the hand is not in the area. This "working area" is basically all the volume above the table.

— **Moving + Area:** the robot looks at the human's hand whenever the hand is in the "working area" and is moving. Else the robot looks at the human head.

In the same way as for the previous scenario, participants were asked to answer several questions for each video. There was 4 5-scales questions for each video concerning the understanding of the human's actions by the robot and the fact that the robot head behavior was adequate, clear and useful to the interaction. Then, the participants were asked to choose which video they preferred between the three videos (with the possibility to select several). These questions can be found in Appendix B. We also let a free space for comments at the end of the questionnaire.



(a) Scores on the questions asked for each video of the first scenario concerning the tracking of human's activity. The score is the addition of the rates (in a scale of 5) of the 4 questions.



(b) Preferences for the first scenario concerning the tracking of human's activity. Numbers represent the number of times where a video was chosen for the preference question (on 59 participants).

Figure 6.3: Results for the the first scenario on the tracking of human's activity. The conditions where the robot looks at the human's hand whenever it is in a "working area " and moving has been significantly preferred to the two others.

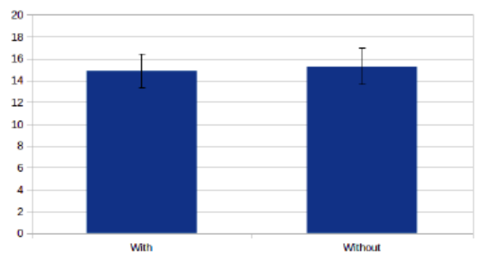
The results of the questionnaire for this scenario can be found in Fig. 6.3. The third condition (where the robot is considering area and movement) has been evaluated significantly higher than the two others ($p < 0.01$). No significant difference has been found between the two others videos. These results give us clues for the construction of a robot head behavior. Indeed, we can see that with a minimal robot behavior (a simple switch between the hand and the head based on the hand position and movement), the robot is able to show its attention and understanding in a relatively acceptable way (rated at 15.78/20 in the 5-scale questions by the participant).

Understanding of the human's action: Then, we asked ourselves if, when an action of the human is detected by the robot, the robot should show its understanding of the action by recording a small stop on the action. We asked participant to watch two videos. These videos were also based on the end of interaction session (the part where the human puts the two last cubes). In each video the human puts the two cubes one after the other and the robot was looking at human's hand and head following the third behavior described previously (where the robot considers area and movement). Then, the different conditions were:

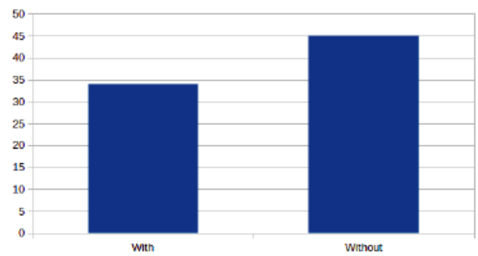
— **With:** in one video the robot was making a stop on the stack each time the human placed a cube on it.

— **Without:** in the other video the robot was not making any stop when the human placed a cube on the stack.

The participants were asked to answer the same questions as for the previous scenario.



(a) Scores on the questions asked for each video of the second scenario concerning the tracking of human's activity. The score is the addition of the rates (in a scale of 5) of the 4 questions.



(b) Preferences for the second scenario concerning the tracking of human's activity. Numbers represent the number of times where a video was chosen for the preference question (on 59 participants).

Figure 6.4: Results for the second scenario on the tracking of human's activity. One condition includes a stop after each action of the human (With) while the other corresponds the the third behavior of the previous scenario (Without).

The results of the questionnaire for this scenario can be found in Fig. 6.4. No significant difference was found between the conditions either concerning the scores in the 5-scales questions ($p=0.33$) or the preferences ($p=0.1093$). Indeed, based on the free comments of the participants, we found that they had difficulties to find the differences between both conditions, and, when they found the differences, they sometime disliked the fact that, when the robot makes a stop (in the condition with as stop), the robot does not follow anymore what the human is doing.

6.4.3 Helping the human to perform his next action

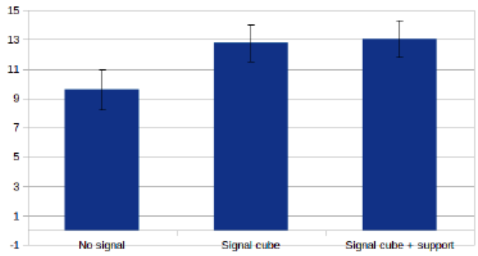
Then, we bear interest to the signals the robot should give with its head concerning the Shared Plan execution. In this part, the videos were based on the middle part of the interaction session (the second cube of the robot and the first of the human).

Human inaction: The first signal we identified as potentially interesting concerns the signaling of an action the human should perform and is not performing. In all tested videos, the robot was putting its second cube on the stack, then, the human was waiting a moment before putting his cube. The different conditions were:

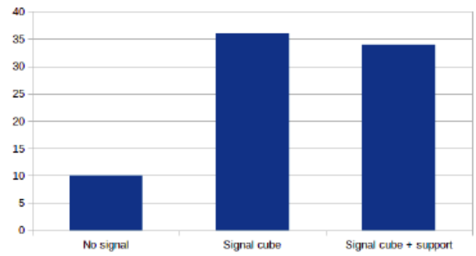
- **No signal:** the robot was not giving any signal to the human, it was just following the "basic behavior" described earlier.
- **Signal cube:** after a small waiting time, the robot gave a signal to the human for him to place his cube on the stack. The signal consisted on looking at the cube to take and then looking back at the human's head.
- **Signal cube + support:** after a small waiting time, the robot gave a signal to the human for him to place his cube on the stack. The signal consisted on looking at the cube, then looking at the stack and finally looking back at the human's head.

In the same way as for the previous scenarios, participants were asked to answer several questions for each video. There was 3 5-scales questions for each video which concerned the fact that the robot head behavior was adequate, clear and useful to the interaction. Then, the participants were asked to choose which video they preferred between the three videos (with the possibility to select several). These questions can be found in Appendix B. We also let a free space for comments at the end of the questionnaire.

The results of the questionnaire for this scenario can be found in Fig. 6.5. The first condition (where there is no signal) has been evaluated significantly lower than the two others ($p < 0.01$). No significant difference has been found between the two different signals given by the robot. These results show us that the studied signal (the robot signals to the human that he should act when he is not acting) is considered important by peoples. One possible explanation concerning the lack of difference between the two ways tested to perform the signal is that, in this task, there is only one place where to put the cube. Indeed, in the signal where the robot looks at the stack, the information given is not so interesting in this configuration. Further investigations can be interesting with scenarios where the human has several choices of actions to execute.



(a) Scores on the questions asked for each video of the first scenario concerning the signaling of human's actions. The score is the addition of the rates (in a scale of 5) of the 3 questions.



(b) Preferences for the first scenario concerning the signaling of human's actions. Numbers represent the number of times where a video was chosen for the preference question (on 59 participants).

Figure 6.5: Results for the first scenario on the signaling of human's actions. In one condition the robot was not giving any signal to the human (No signal). In another conditions the robot was giving a signal consisting of looking at the cube to place and looking back at the human's head (Signal cube). In the last condition the robot was giving a signal consisting of looking at the cube to place, the stack and then looking back at the human's head (Signal cube + stack).

Turn-taking: Then, we investigated the use of a signal to help turn-taking. We specifically focused on the signal needed at the end of a robot action if the next action to be performed is a human action. In this scenario, we compared 4 videos where the robot puts its second cube and then the human puts his first cube. The different conditions were:

— **Signal - Before:** the robot was giving a signal to the human at the end of its action. This signal consisted in looking at the human's head, looking at the cube he should put on the stack and then looking back at the human's head. The robot was giving the signal while retracting from its action.

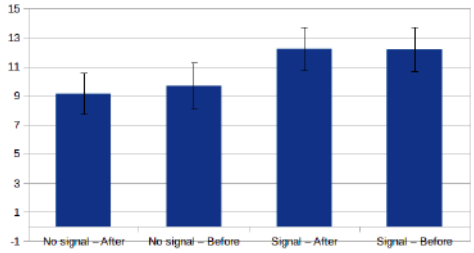
— **Signal - After:** the robot was giving a signal to the human at the end of its action. This signal consisted in looking at the human's head, looking at the cube he should put on the stack and then looking back at the human's head. The robot was giving the signal after retracting from its action.

— **No signal - Before:** the robot was not giving signal to the human. Its looks at the human's head as soon as its action is over(while retracting).

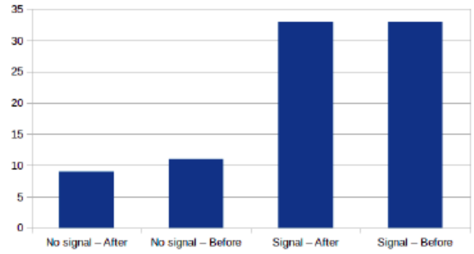
— **No signal - Before:** the robot was not giving signal to the human. Its looks at the human's head only after retracting from its action.

The participants were asked to answer the same questions as for the previous scenario.

The results of the questionnaire for this scenario can be found in Fig. 6.6. The two conditions where there is a signal have been evaluated significantly higher than the two others ($p < 0.01$). No significant difference has been found between the two conditions with a signal as well as for the two conditions without a signal. These results show the usefulness of a signal at the end of robot actions in order to help



(a) Scores on the questions asked for each video of the second scenario concerning the signaling of human's actions. The score is the addition of the rates (in a scale of 5) of the 3 questions.



(b) Preferences for the second scenario concerning the signaling of human's actions. Numbers represent the number of times where a video was chosen for the preference question (on 59 participants).

Figure 6.6: Results for the second scenario on the signaling of human's actions. In one condition the robot was not giving any signal to the human and was looking at his head after retracting (No signal - After). In another condition, the robot was not giving any signal to the human and was looking at his head while retracting (No signal - Before). In the two others conditions the robot was giving a signal consisting in looking at the the human's head, looking at the cube to place on the stack and looking back at the human's head. In one condition the robot was giving the signal after retracting from its action (Signal - After) and in the other while retracting (Signal - Before).

turn-taking. Apparently the timing of the signal (during or after retracting) has not been found important by the users. However, this result can be questioned as there was no question concerning the perception of the efficiency of the interaction. Indeed, giving the signal while the robot is retracting allows to gain time and fluidity but this was tiny and most users did not even noticed it, focusing on the fact that the robot is producing a signal or not.

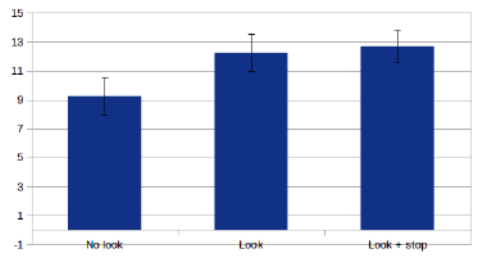
6.4.4 Finding the priority target

In the last scenario of the study, we focused on the decision to take when the robot has several possible targets to look at. We studied here the case where the robot is executing an action (and so has a target for its action) and the human performs another action at the same time (and so the robot should also look at the human's action). In the videos, the robot is still placing its second cube on the stack when the human picks his next cube (in prevision of placing it on the stack). In each video, the robot looks at the stack when putting its cube. Then, the different conditions were:

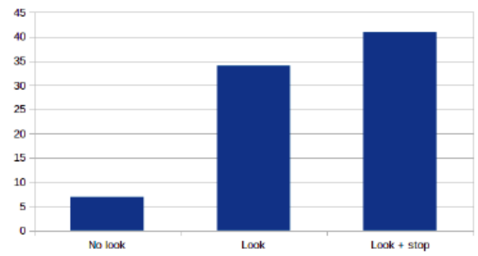
- **No look:** the robot does not look at all at the human action.
- **Look:** the robot looks at the human action (it looks at the human's hand and goes back to the stack) when it is executed without interrupting its action.
- **Look + stop:** the robot stops its own action to look at the human action

and then restarts its action.

In the same way as for the previous scenarios, participants were asked to answer several questions for each video. There was 3 5-scales questions for each video concerning the fact that the robot behavior was adequate, clear and useful to the interaction. Then, the participants were asked to choose which video they preferred between the three videos (with the possibility to select several). These questions can be found in Appendix B. We also let a free space for comments at the end of the questionnaire.



(a) Scores on the questions asked for each video of the scenario concerning the priority target. The score is the addition of the rates (in a scale of 5) of the 3 questions.



(b) Preferences for the scenario concerning the priority target. Numbers represent the number of times where a video was chosen for the preference question (on 59 participants).

Figure 6.7: Results for the scenario concerning the priority target. In one condition the robot was not looking at the human's action (No look). In another condition the robot was looking at the human's action without interrupting its action (Look). In the last condition the robot was interrupting its action to look at the human's action (Look + stop).

The results of the questionnaire for this scenario can be found in Fig. 6.7. The condition where the robot does not look at the human's action has been evaluated significantly lower than the two others ($p < 0.01$). No significant difference has been found between the two conditions where the robot looks at the human's action. These results show us that even if the robot is busy doing something else it is important that it looks at human's action to acknowledge the fact that it perceived the action. If the robot action requires the full focus of the robot head, the robot can interrupt briefly its action in order to look at the human's action without degrading the acceptability of its behavior.

6.5 The robot head behavior

Based on our previous analysis and the performed on-line user study we made a proposal for an architecture to generate a robot head behavior. This behavior is for now computed based on a Joint Action with only one human.

The architecture can be found in Fig. D.18. We can find in this architecture the four "families" described in Sec. 6.3 (human observation, robot action, dialogue and coordination). Based on informations concerning humans, the robot constantly computes a target (either an object or a part of the human body) to look at. In addition, when the robot is performing an action, it also computes a target for the action. In a same way, when the robot is speaking with a human, another target is computed. Based on the current state of the Shared Plan, the robot also computes signals to give to the human. Finally, based on the different targets and signals, an arbitration module chooses at each time the final target to look at.

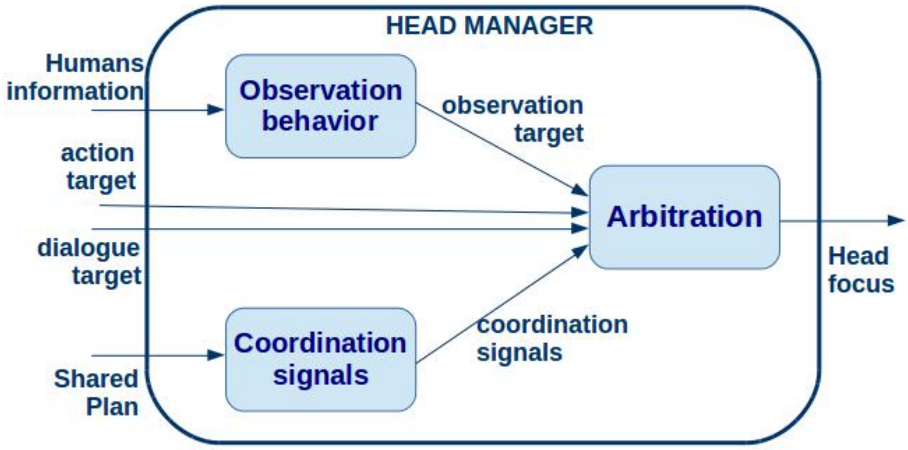


Figure 6.8: The architecture to generate the robot head behavior. The arbitration modules chooses where the robot should look based on several possible behaviors and signals. The possible behaviors concern the human activity, the current robot action and the current state of the dialogue. Coordination signals are computed based on the current Shared Plan.

In this architecture, a target is described as a point to look at (either an object or a part of the human body) and an associated priority.

$$target =< point, priority >$$

A signal is described as an ordered set of points to look at with associated durations, a priority, an expiration time and possibles expiration conditions (events).

$$signal =< points, durations, priority, expiration\_time, expiration\_conditions >$$

This architecture is still a work in progress, there are still needs of deepening for some parts and the parametrization of the priorities and durations are still needed.

6.5.1 Observation target

We saw previously that the robot should show its interest and understanding of the human activity. We studied in the previous section how we can show the interest of the robot in the human's actions based on the position and movements of his hand. We keep the condition of the study which was the most highly marked: the robot looks at the human's hand whenever the hand is moving (with a small hysteresis in order to avoid going back and forth between the hand and the head) and into a "working area". Working areas can be defined dynamically and attached to objects (e.g. everything above a table).

Additionally to this behavior, another important feature is joint attention. When the human stares at an object for a sufficient time, the robot should also look at this object. In a same way, if the human stares at the robot, the robot should return his look. The given algorithm to compute the point corresponding to the human observation can be found in Alg. 9.

Algorithm 9 Computation of the point to look based on human's activity

Require: human's hand position and movement, objects position, human's head position and target
  if human's head target = robot > t **then**
    $point = human's\ head$ ▷ *the human stares at the robot*
  else if human's head target = object > t **then**
    $point = object$ ▷ *the human stares at an object*
  else if human's hand is moving & human's hand is in "working area" **then**
    $point = human's\ hand$ ▷ *tracking human's action*
  else
    $point = human's\ head$ ▷ *by default, looking at human's head*
  end if
  return *point*

This behavior is for now really basic. One of the reason is that we want a behavior which can work with few information concerning the human (here only concerning the head and one hand). Moreover, we strongly think, even if it remains to be proved, that a minimal behavior is sometime preferred to a more complex one as soon as all the needed components are here.

6.5.2 Robot action and dialogue targets

We saw the importance of having a head behavior in adequacy with the robot current action. Consequently, the action target is computed directly by the component in charge of executing robot actions and depends of the action executed. The point to look at is in most cases the "target" object (e.g. the object to pick or the support where to place). The priority associated with this point principally depends of the fact that the action requires the head focus (e.g. at the beginning of a pick action, the priority will be higher to ensure the good perception of the object). The module in charge of executing actions has also access to the point the robot is looking at in order to adapt the action execution (e.g suspend if necessary).

In a same way, the target associated to dialogue is directly computed by the module in charge of the dialogue. It mainly consists in looking at the human when talking and at the objects the robot or the human is referring to.

6.5.3 Coordination signals

We saw the importance of coordination signals during Joint Action execution. In this work, these signals are based on the execution of the Shared Plan. The first two signals are the ones studied in the on-line study. The first one is computed whenever a human has an action to execute and is not performing it after a certain amount of time (and that he is not doing another action). In this case, the robot signals to the human the action he should execute. This signal consists of looking at the human's head, then the different points of interest of the action to execute and finally looking at the human's head again. The points of interest of an action are ordered and defined for each high level action (e.g. for the pick and place action it is the object to pick and, then, the support where to place). This signal should have a relatively long expiration time as its timing is not crucial. It should also have the execution of the awaited action in its expiration conditions.

Another signal concerns the help of turn-taking. After each robot action, if this action unlocks another action attributed to the human (there is a causal link between the two actions), the robot sends a signal to the human to execute the action. This signal consists of looking at the human's head, then the different points of interest of the action to execute and finally looking at the human's head again. This signal should have a small expiration time as it has to be performed just after the robot action. It should also have the execution of the awaited action in its expiration conditions.

Finally, we saw the importance of being attentive to a human's actions, even if the robot is currently busy doing something else. To do so, when the human performs an action, a signal is computed with a high priority as well as a small expiration time because the signal does not make sense if it comes later after the action. This signal consists of looking a the point of interest of the executed action.

6.5.4 Arbitration

The arbitration of the different points to look at and the different signals to send is based on priorities. The arbitration module first chooses the point to look at with the higher priority between the ones from human observation, dialogue and robot action. Then, when a signal is sent by the coordination module, it is stored in a waiting list (ordered by priorities). If the priority of a signal in the waiting list is higher than the current priority, the robot executes the signal. In order to keep signals consistent, a signal started can not be interrupted even if there is another signal or point to look at with a higher priority. When a signal expires (the expiration time is reached or an expiration condition is true) the signal is removed from the waiting list.

There may be some missing signals in the current version of the architecture. However, we strongly think that it has been done in a sufficiency generic way in order to integrate new signals if needed.

6.6 Conclusion

In this chapter, we studied non-verbal behavior during human-robot Joint Action, and more especially the robot head behavior. A first bibliographic study has been done in order to identify the needed components of a robot head behavior adapted to the Joint Action. Then, some specific parts of these components have been studied with an on-line video based study. Finally, an architecture to generate the robot head behavior has been proposed. This architecture, still in construction, should allow the robot to produce a head behavior which provides the needed information in order for the robot to show its attention, make its action understandable, dialogue and coordinate the Joint Activity based on the Shared Plan.

Chapter 7

Combining learning and planning

Contents

7.1 Motivation

When it comes to decisional process in robotics, two main schools of thought can be distinguished: machine learning and deterministic processes such as planning or states machines. Both ways have their advantages and disadvantages. Learning is usually "cheap" (the decision process is quick) and always proposes a solution to a given problem. However, learning requires either a big amount of data or a long period of learning. Moreover, during the learning period, the robot can produce inconsistent behavior which can be confusing for a potential human collaborator. On the other hand, planning can take into account humans through social rules and ensure the validity of a whole solution. However, planning does not learn from human behavior, and, when it comes to complex tasks or environments, it can become slow to propose a solution. The idea of this work is to propose a solution where we combine planning and learning in the context of human-robot interaction in order to take advantage of both.

This work has been done in collaboration with ISIR at Paris and more particularly with another PhD student Erwan Renaudo. It has been done in the context of the RoboErgoSum ANR project [1]. This work is based on the work of [Renaudo 2014] and has been the subject of a publication in a workshop at the RoMan conference [Renaudo 2015] as well as a part of a journal article [Khamassi 2016].

7.2 Background

7.2.1 Inspiration from neurosciences

Several works on living being behaviors have been done in the late 19th century - beginning of the 20th century - with experiments on mammals. One pioneer work concerning the learning process is the experiment of the cat in a box [Thorndike 1911]. In this experiment, a cat is put in a box each time it is hungry. The cat can see food outside of the box and a system of lever allows it to open the box. Each time the cat is put in the box it takes less time to go out. This experiment allows to show the principle of learning through trial and error.

Latter, studies have highlighted two main kinds of behavior during decision-making tasks. **Goal-directed behaviors** are governed by estimates of action-outcome contingencies (i.e. decision-making relies on the prior estimation of the outcome expected after an action or an action sequence) and are mainly active at the beginning of the task. Then, when the animal is well trained in the task under stable conditions, a transfer of control to **habitual behaviors** governed by stimulus-response associations occurs [Dickinson 1985]. When rodents, monkeys or humans start a new decision-making task, they appear to initially rely on their goal-directed system. They take time to analyse the structure of the task in order to build an internal model of it, and make slow decisions by planning and inferring the long-term consequences of their possible actions before deciding what to do next. Then as their performance gradually improve, they appear to make quicker and quicker decisions, relying on their habitual system which slowly acquires simple stimulus-response associations to solve the task. Finally, when subjects restart to make errors after a task change, they appear to restart planning within their internal model and thus slow down their decision process before acquiring the new task contingencies [Balleine 2010, Dolan 2013]. The coordination of these two learning systems allows mammals to avoid long and costly computations when the environment is sufficiently stable, while still enabling animals to detect environmental changes requiring to update their internal model and replan.

In computational neuroscience models, these behaviors are modeled using the theory of Reinforcement Learning [Sutton 1998]: model-based and model-

1. http://roboergosum.isir.upmc.fr/

free algorithms provide a direct analogy with goal-directed and habitual behaviors [Daw 2005]. More recently, different computational criteria have been proposed to decide when to shift between model-based and model-free experts [Pezzulo 2013, Lesaint 2014, Viejo 2015]. Applied to neuroscience tasks, the work from [Daw 2005] proposes that the most certain expert gets control on the agent, while [Keramati 2011] balance speed and accuracy using the cost of planning versus the gain of information. A third approach proposes, in the context of navigation strategies, that a coordination module learns by reinforcement the most efficient behavior (in terms of average obtained reward) in each state [Dollé 2010].

7.2.2 Learning in human-robot interaction

A major part of robotics decision-making algorithms are based on planning processes which take into account a great number of information ([Ingrand 2014]). These approaches to decision-making could be seen as similar to what neuroscientists call the goal-directed system, except that there is most of the time no learning in the system. Such approaches have been extended to HRI by taking into account human-aware costs such as social-rules and humans comfort and preferences [Cirillo 2010, Lallement 2014].

Besides, robots learning abilities are still very limited and require the injection of important prior knowledge by the human in the robot's decision-making system. Early applications of reinforcement learning (RL) algorithms to robotics [Hayes 1994, Morimoto 2001, Smart 2002] - some of which being neuro-inspired - produced limited progresses, due to applications to relatively simple problems (with a small number of states and actions), to slowness in learning and to systematic instability observed throughout the learning process. More recent applications of RL to robotics have permitted to deal with more complex and continuous action spaces, enabling to learn efficient sensorimotor primitives [Kober 2011, Martins 2010, Stulp 2013]. These approaches have been extended in HRI to allow robots to learn to collaborate with humans. In several works, the reward signal is interactively assigned by the human [Kaplan 2002, Knox 2012] while other works use the human to provide demonstrations to the robot [Nicolescu 2003, Thomaz 2006]. A method of cross-training is used and compared to standard reinforcement learning algorithms in the context of human-robot teamwork in [Nikolaidis 2013]. Cross-training is an interactive planning method in which a human and a robot iteratively switch roles to learn a shared plan for a collaborative task. Such approaches to decision-making could be seen as similar to what neuroscientists call habitual behaviors.

Even if we can find more and more interesting works in HRI concerning planning and learning for the robot to collaborate with humans, there is no work to our knowledge concerning how to combine both approaches into a robotics architecture.

7.3 Experts presentation

Inspired from neuroscience theories and based on the previous work of [Renaudo 2014] the aim of this work is to combine goal-directed and habitual behaviors in the context of human-robot Joint Action. To do so, we use two experts which implement these two kinds of behavior. The goal-directed behavior is produced here by HATP (Human-Aware Task Planer), a task planner which has proved its efficiency in the field of human-robot interaction. A Qlearning algorithm allows to implement the habitual behavior. We will describe in this section these two experts and their respective strengths and weaknesses. The next sections will show how we combined those two experts into two different architectures.

7.3.1 HATP

In our work, the goal-directed behavior is provided by HATP, an HTN (Hierarchical Task Network, [Erol 1994]) task planner which has been conceived to work in the context of human-robot collaboration. As a HTN planner, HATP uses known preconditions and effects of actions in order to find the best plan that reaches the given goal. It takes as input a list of all possible actions and their description in terms of preconditions and effects and also a description of the current world state as a set of predicates. Then, it looks for the combination of actions that minimizes the solution cost. This cost is computed based on execution time and human-aware costs (e.g the balance of efforts between agents or the waiting time of the human partner). This plan is meant to be executed step by step until the goal is reached. An example of such a plan can be found Fig. D.19.

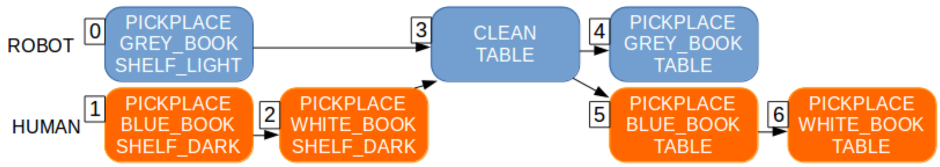


Figure 7.1: An example of a Shared Plan computed by HATP. This plan allows a human and a robot to clean a table by removing all objects on it, cleaning it and then putting back all previous objects.

7.3.2 Qlearning algorithm (MF)

The habitual behavior is provided by a model-free reinforcement learning algorithm (MF) that directly learns the state–action associations by caching in each state the earned rewards in the value of each action. In this implementation, the algorithm is implemented as a neural network (see Fig. D.20). The network input neurons represent the different possible states and the output neurons encode the estimation of action values in the current state. The weights are modified to associate each state with the most rewarding action in the current task.

A method similar to [Brafman 2002] is used to compute the value $Qt(St, a_j)$ of an action a_j in a certain state S. $Qt(St, a_j)$ is represented as the scalar product between the input vector and the weights $W_j = (w_{0j}, ..., w_{Nj})$ linking to this action:

$$Qt(St, a_j) = W_j^t \ . \ (S_t, 1)$$

here we set weights at a positive value to provide an initial optimistic estimate of action values ($w_0 = 0.5$). Weights W_t are updated according to the Qlearning algorithm [Watkins 1989]. The Reward Prediction Error δ is spread over the weights of the previously active input and the action a done in the corresponding state:

$$\delta = r_t + \gamma_{Hab} \ . \ \underset{b \in A}{max}(W_b^{t-1}) \ . \ S_t - (W_a^{t-1} \ . \ S_{t-1})$$

$$W_a^t = W_a^{t-1} + \alpha_{Hab} \ . \ \delta / \sum_n s_n$$

with r_t the instant reward received for performing a in S_{t-1} , α_{Hab} the learning rate, γ_{Hab} the decay rate of future rewards. The weights are updated locally: only the state from which the action has been performed is updated. Thus, it requires for the agent to visit every known state of the problem to update values.

7.3.3 Experts comparison

The two different experts have really different ways to decide of the next action to execute. Both methods have their advantages and disadvantages:

— HATP looks for a complete solution to achieve the given goal while the MF only looks for the next action which maximizes the probability to get a reward. Consequently, HATP ensures the feasibility of the solution proposed but could find itself in a state where it does not find a valid solution and so where it will not be able to propose an action. In the other hand, the MF does not ensure that its proposed action allows to achieve the goal but will always propose an action to perform.

— As HATP computes a whole plan to achieve the goal, its cost, in the sense of time to take a decision, is far bigger than the one of the MF which only proposes the next action. However, this difference needs to be weighed by the fact that as an HATP plan is composed of several actions, this cost is not needed at each step of the task. Moreover, this cost stay acceptable in a not so complex task.

— HATP is conceived to produce a robot behavior understandable and acceptable by the human. The actions it proposes will produce a consistent behavior of the robot with which one the human can easily collaborate. For its part, the MF has a long period of learning during which one the behavior produced is inconsistent and can be really disturbing for a human collaborating with the robot. Moreover, each time a change happens in the task, a new learning phase is needed. However, the MF is able to learn to adapt its behavior to the human whereas the HATP policy is defined off-line and can

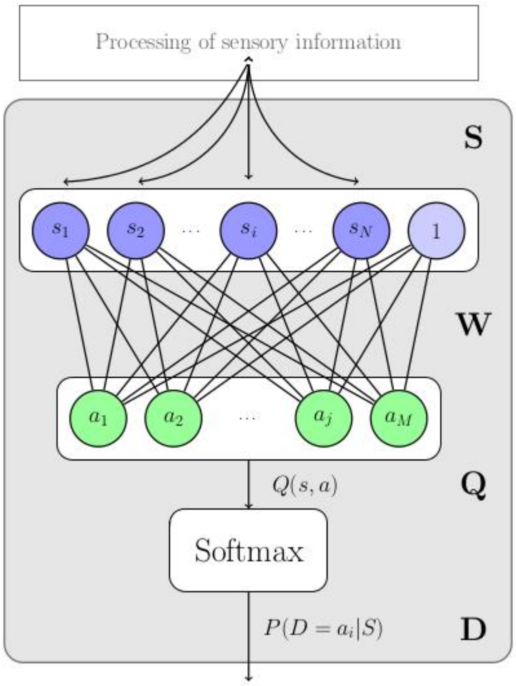


Figure 7.2: Habitual expert, modeled as a Qlearning algorithm implemented as a neural network. The expert receives a state S which is projected onto the input neurons s_i, defining an input activity. This activity is propagated through the network weights W to generate activity of the action layer. This activity corresponds to the values $Q(S, a_j)$, with each neuron coding for a distinct action. This value distribution is converted in probability distribution using a softmax function, which allows the expert to make a decision D on the next action to perform.

not be updated with the behavior of the human during the interaction.

7.4 First architecture: a proof of concept

7.4.1 Control architecture

The first architecture in which we tried to combine the two experts is the one illustrated in Fig. D.21. In this architecture the two experts are placed in parallel. The execution of a task by the architecture follows several steps:

— The Situation Assessment module receives data from perception and maintains the world state representation. This world state is represented with predicates (see Sec. 7.4.2).

— The supervisor uses the current world state to compute the reward sent to the MF. This reward is a boolean which is true if the current goal is achieved (see Sec. 7.4.2). The supervisor also sends to the MF the last tried action (which was not necessarily the one proposed by the MF) in order to update

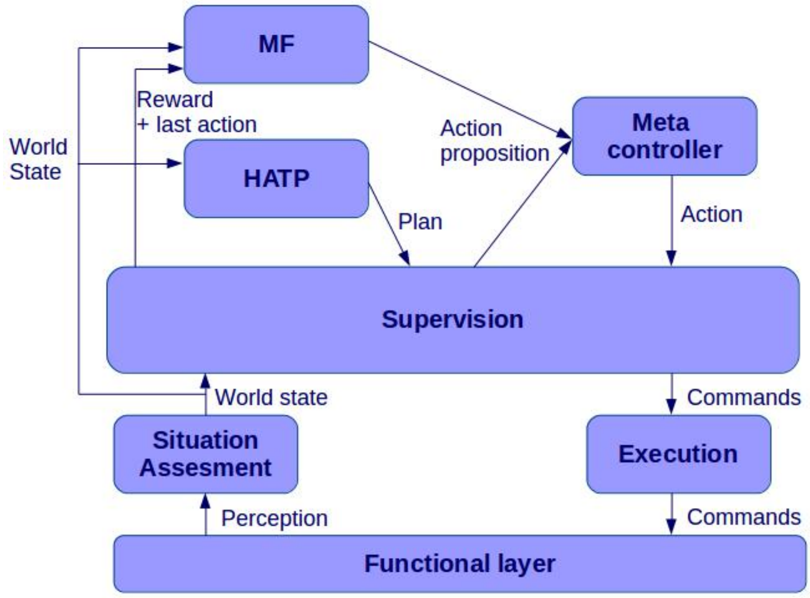


Figure 7.3: First tried architecture to combine the two experts. The Situation Assessment module gets data from perception and maintains the current world state. This world state is used by the supervision to compute the reward and by the experts to take decisions. The propositions of the two experts are sent to the meta controller which decides of the action to execute. The supervisor executes the action with the help of lower execution modules.

the learning.

— The experts decide of the next action to execute based on the current world state. The action proposed by the MF for a given world state is sent directly to the meta controller. Concerning HATP, the supervisor monitors the execution of its plan and sends the next action to execute to the meta controller. A new plan is computed by HATP at the beginning of the task or whenever an unexpected situation happens (an action from the plan fails, the human executes an unexpected action or the robot executes an action proposed by the MF which is not in the current plan).

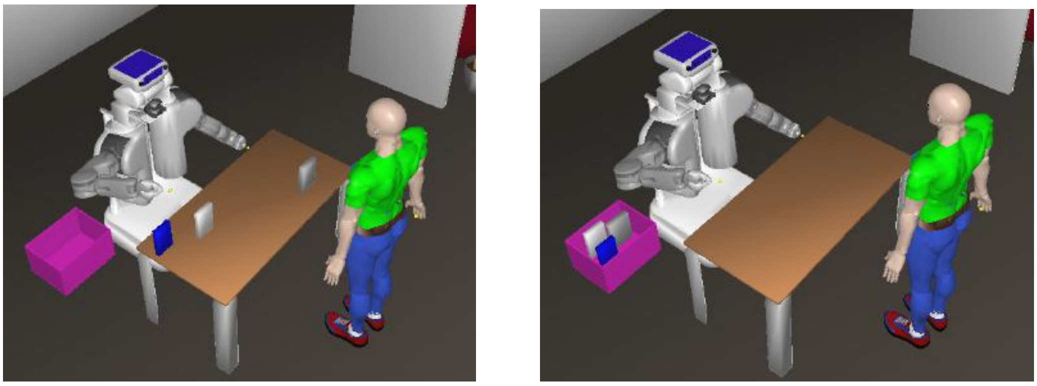
— Once the proposition of action from each expert is received, the meta controller decides which action the robot should execute. In this first implementation the meta controller uses a random arbitration: the action is chosen with an equal probability for each expert.

— The supervisor executes the chosen action with the help of lower execution modules (motion planning, control, ...).

These steps are executed one by one until the goal is achieved.

7.4.2 Task

This first architecture has been tried in a simple task. Moreover, as the learning part of the architecture requires long learning periods, the tests have been done in simulation.



(a) Initial situation (b) Final situation

Figure 7.4: Description of the task used with the first architecture. In this task, the human and the robot have to remove all the objects of the table and put them in the pink box. At the beginning of the interaction two objects are accessible only by the robot and another one only by the human. The box is accessible only by the robot.

In the chosen task a human and a robot have to "clean a table" together. To do so, they need to remove all the objects from the table and put them in a box (see Fig. D.22). At the beginning of the interaction two objects are accessible only by the robot and another one only by the human. The box is accessible only by the robot. To achieve the goal, several actions can be executed by the agents:

- **Pick an object:** both agents can pick an object accessible by them.
- **Store an object:** the robot can store an object it has in hand in the box near itself.
- **Give an object:** the robot can give an object to the human.
- **Take an object:** the robot can receive an object from the human
- **Wait:** the robot can wait for the human to execute an action.

All these actions have an impact into the world state. This world state is estimated by the Situation Assessment module and represented with predicates which can be either true or false. For this task, we consider the following predicates:

- **<Object, isReachableBy, Agent>:** these predicates represent for each object if it is reachable by the human or the robot.
- **<Object, isIn, Box>:** these predicates represent the fact that an object has been stored in the box.
- **<Agent, hasInHand, Object>**: these predicates represent the fact that the human or the robot holds an object.

These predicates allow the experts to take their decisions but also the supervisor to compute the reward needed by the MF. The robot will receive a reward whenever all objects are in the box and it performs the *Wait* action. We chose to impose to the robot to perform a *Wait* action at the end of the task in order for it to learn that the task is over and that no more action is needed.

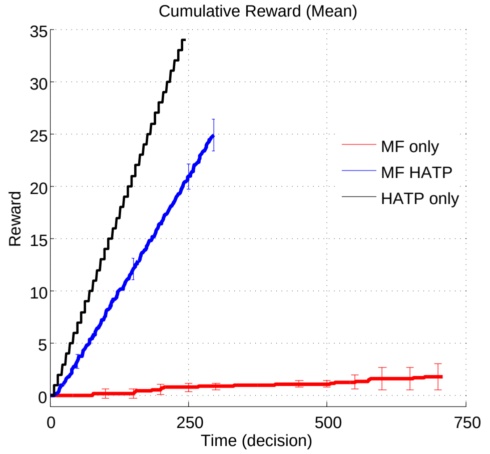
To test our architecture, we compare its performances to the performances of the system running with only the MF and only HATP. We run the experiment in all conditions with a fixed time limit. At the beginning of an experiment the set-up was put at the initial situation (Fig. D.22(a)). Once the task is achieved and a reward is obtained by the robot, the set-up is put back to the initial situation and the task can be performed again.

As we run the task in simulation, the behavior of the human is also simulated. We chose here to have a collaborative human: it performs all actions HATP planned for him and participates to handover whenever the robot requires one.

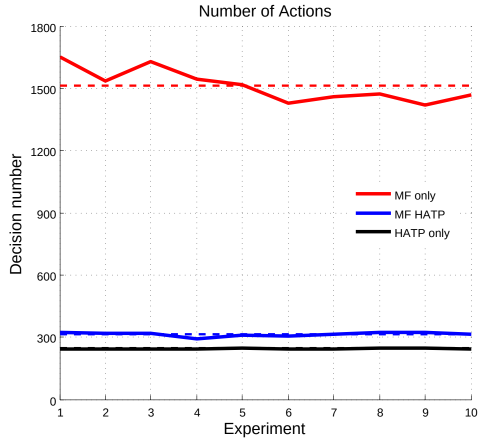
7.4.3 Results

The main criteria used to evaluate our system is the cumulative reward obtained in each run (i.e. the number of time the human and the robot manage to achieve the task in a fixed amount of time). We run 10 times the experiment in each condition (MF only, HATP only and the combination of both) for a duration of approximately 30 minutes. The number of rewards obtained are presented in Fig. D.23(a). All experiments last the same fixed time, but the number of decisions taken at the end may vary. We observe a poor performance of the MF alone, which is not able to solve the task more than three times. As the MF has no initial knowledge, it has to discover the right sequence of actions, which is non trivial with the given number of possible states and actions. The random combination HATP-MF is performing much better than the MF alone, solving the task 25 times in average. However, HATP alone performs even better solving the task 34 times in average. Indeed, the task is easy enough to solve for HATP and the time required to find a plan is negligible here. As the simulated human always performs the actions planned by HATP, the plan found by HATP is always optimal and will never change during the task execution. Accordingly, the random combination of HATP and the MF performs worse as it can include actions proposed by the MF that make the plan non optimal.

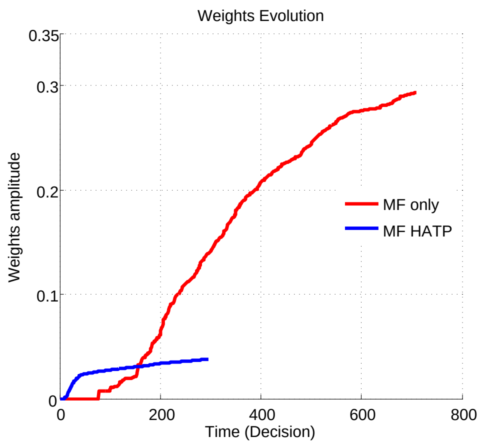
Fig. 7.5(b) shows the number of actions proposed to the supervision system during each experiment. We can see that the MF alone suggests twice to three times more actions than HATP or the combination in the same given time. This is mainly due to the way each Expert decides: the MF only needs to compute the values of each action (which is propagating the state activity to action neurons) and to draw an action from the resulting probability distribution. It proposes a lot of unfeasible actions and the supervision system will not spend time to execute them as it will stop to the preconditions verification. HATP checks for action preconditions when planning and so, for each of the action proposed by HATP, the supervisor spend time to execute it (or try to execute it if the action is not really feasible



(a) Mean cumulative reward on 10 simulations where the robot repeatedly fulfils the task. Errorbars represent the standard deviation from the mean every 100 decisions.



(b) Number of actions tried per experiment. Dashed line is the mean number of action depending on the control method (MF only, HATP only or combination)



(c) Mean MF connection weights evolution for MF alone and MF and HATP combination. The amplitude is defined as the sum of the absolute value of weights. Weights are initialized to zero, thus the higher the amplitude is, the more the MF has learnt which action to do.

Figure 7.5: Performance of the developed system compared to systems with only the MF and only HATP. The results are for 10 runs of approximately 30 minutes in each condition.

according to the geometry). The number of actions suggested by the combination of Experts is closer to the one with HATP alone while remaining lightly higher. It can be explained by the fact that a part of the actions proposed by the combination comes from HATP and, for the ones coming from the MF, HATP helps it to learn a solution faster, causing it to propose less unfeasible actions.

Finally, we analyze the effect of combination on learning of MF in Fig. D.23(b).

Learning is evaluated by weights amplitude, namely the sum of weights absolute value over actions. The MF starts with weights initialized to zero, each learning step increases or decreases the value of some of the weights, until convergence. The figure shows that learning occurs much earlier for the combination of Experts than when the MF is alone. The combination has a bootstrapping effect and the knowledge about the task from HATP is transferred to the MF. This shows that a human-provided a priori knowledge can be used to guide exploration and learn quicker. Even if not tested in this experiment, this means that a change in task condition for which HATP can find a new plan can be learnt quickly by the MF, so the robot will be able to adapt to the new conditions without taking too much time.

7.4.4 Intermediate conclusion

The first results obtained with this architecture allow to show that the combination of HATP and the MF allows to bootstrap the MF and to learn faster a policy to achieve the goal.

However, this task is too simple for HATP to be in difficulty when deciding alone. The purpose of the second task and architecture presented in the next section is to show the benefits of the combination of the two experts and more particularly how HATP can benefits from the MF. Moreover, we want to test the reaction of our system to changes in the task as well as a more elaborated arbitration criteria for the meta controller.

7.5 Second architecture: the limitations

7.5.1 Control architecture

One of the advantage of the MF against HATP is its computation time. In the previous architecture, both experts where consulted before the meta controller took a decision. Consequently, even if the MF was chosen, we still lose time to compute plans with HATP. In order to solve this issue, we modified the previous architecture as shown in Fig. D.25.

In the new architecture, the meta controller is placed upstream from the two experts. Consequently, the order of the previous steps during a task is also slightly modified:

- The Situation Assessment module still receives data from perception and maintains the world state representation.
- The supervisor sends the needed data concerning both experts to the meta controller in order for it to take a decision (see below).
- Once the meta controller decision taken, we look for the action proposed by the selected expert. If the MF is chosen it directly sends its action to the supervisor as well as data concerning its decision (see below). If HATP is chosen, if needed, the supervisor asks for a new plan, else it directly executes

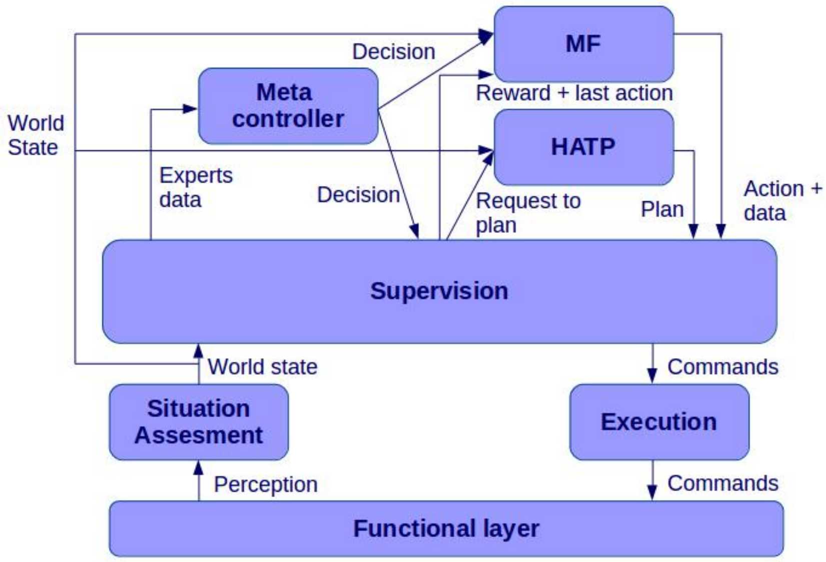


Figure 7.6: Second tried architecture to combine the two experts. The Situation Assessment module gets data from perception and maintains the current world state. This world state is used by the supervision to compute the reward and by the experts to take decisions. Here the meta controller is placed upstream from the two experts. It first decides which expert should propose an action. Then, the supervisor executes the action of the chosen expert with the help of lower execution modules.

the next action of the current plan. A new plan is needed at the beginning of the task or whenever an unexpected situation happen (an action from the plan fails, the human executes an unexpected action or the robot executes an action proposed by the MF which is not in the current plan).

— The supervisor still executes the chosen action with the help of lower execution modules (motion planning, control, ...).

In this architecture, we also introduced a new arbitration criteria for the meta controller. This criteria is based on the cost of each expert (duration to find a solution) and its prediction error. For the MF, the prediction error is the difference between the probability for the proposed action to lead to a reward and the actual received reward. For HATP, the prediction error is 0 if after the execution of the proposed action the world state corresponds to what HATP predicted (based on the action effects) and 1 if it differs.

$$P_t^E = \alpha \ . \ err_t^E + \beta \ . \ cost_t^E$$

with P_t^E the probability for the expert E to be chosen by the meta controller at a time t, err_t^E the prediction error of the expert E at a time t and $cost_t^E$ the cost of an expert E at a time t. α and β are parameters. The prediction error and the cost of the experts are averaged through time:

$$err_t^E = (1 - \gamma_{err}) \ . \ err_{t-1}^E + \gamma_{err} \ . \ err_t^E$$

$$cost_t^E = (1 - \gamma_{cost}) \ . \ cost_{t-1}^E + \gamma_{cost} \ . \ cost_t^E$$

with γ_{err} and γ_{cost} parameters.

7.5.2 Task

The previous task was too simple to have difficulties with HATP as the only expert. The new task is an upgrade of the previous one with several additions.

More complex task A first way to complexify the task for HATP is to increase the combinatory of the task. Indeed, there was not too much ways to solve the previous task, so, HATP didn't need too much time to compute a plan. The goal of the new task is still to "clean a table", however, there are now two different boxes where to put the objects. The blue objects have to go in the blue box and the green objects have to go in the green box. We increased the number of objects in the task: at the beginning of the interaction 6 objects (3 blue and 3 green) are randomly placed on 7 possible placement in the table (see Fig. 7.7(a) and Fig. 7.7(b)).

We also add some new possible actions for the robot:

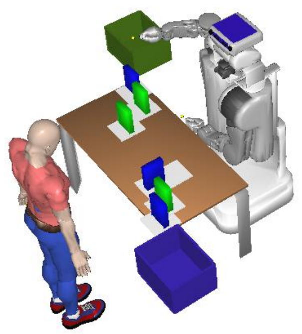
- **Pick an object:** both agents can still pick the objects accessible by them.
- **Store an object:** the robot can store an object it has in hand in a box of the same color accessible by itself. The human can store an object it has in hand in the blue box.
- **Give an object:** the robot can still give an object to the human.
- **Take an object:** the robot can still receive an object from the human
- **Place an object on a placement:** both agents can place an object they have in hand on a placement accessible by themselves.
- **Navigate to another position:** the robot can navigate to another position in order to change the objects it can reach. The two possible positions for the robot are the one in Fig. 7.7(a) and Fig. 7.7(b)) and the one in Fig. 7.7(c).
- **Wait:** the robot can still wait for the human to execute an action.

The predicates used to represent the world state also changed. They are now composed of:

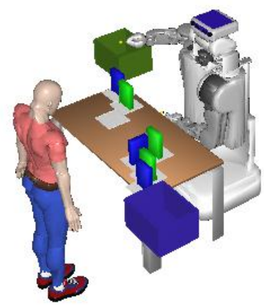
- **<Object, isReachableBy, Agent>:** these predicates represent for each objects if they are reachable by the human or the robot.
- **<Placement, isReachableBy, Agent>:** these predicates represent for each placement if they are reachable by the human or the robot.
- **<Box, isReachableBy, Agent>:** these predicates represent for each box if they are reachable by the human or the robot.
- **<Object, isIn, Box>:** these predicates represent the fact that an object has been stored in a box.

— **<Agent, hasInHand, Object>**: these predicates represent the fact that the human or the robot holds an object.
— **<Object, isOn, Placement>**: these predicates represent the fact that an object is on a specific placement.
— **<Robot, isAt, Position>**: these predicates represent the position of the robot (Position are the two possible places it can navigate to).

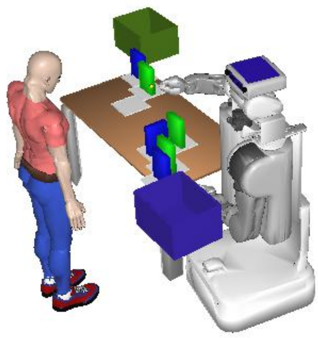
In this task, a reward is given to the robot whenever all objects are in a box.



(a) One possible initial set-up. In this situation the robot can access four objects (two blue and two green) as well as the green box.



(b) Another possible initial set-up. In this situation the robot thinks it can access the blue object in the middle of the table. However, the green object in front of it blocks its access.



(c) One possible way for the robot to access the blue object it was not able to reach is to move to another position.

Figure 7.7: Description of the task used with the second architecture. In this task, the human and the robot have to remove all the objects of the table and put them in the box of the same color. At the beginning of the interaction several objects are accessible by the robot, others by the human and others by both agents. The green box is accessible by the robot and the blue one by the human. The placements are the white squares on the table.

As there are more objects and more actions to perform for the robot, the number of possible ways to achieve the task highly increases. Indeed, in the initial set-up of the previous task HATP needed around 145ms to find a plan to achieve the goal. In this new task it takes around 22s to find a plan. Consequently, the difference of cost between the MF and HATP should make more difference in this experiment.

Difference between planning and geometry In order to get closer from possible real life situation, we introduced a geometrical problem in the task. Indeed, sometimes it can happen that the knowledge computed by the robot is not accurate and that, consequently, the computed plan is not valid at execution. In our task, there are two placements in the middle of the table (accessible both by the human and the robot) which are close to each other. Each time an object is on one of these placement, the robot thinks it can reach it. When there is an object in only one of the placement (as in Fig. 7.7(a)) the robot can effectively reach the object. However, when there is an object in both placements, the robot cannot reach the one in the farthest placement as the other one blocks its access (see Fig. 7.7(b)). The Situation Assessment is not able to differentiate the two situations and in each case it will estimate that all objects in these placements are reachable by the robot. The robot will discover that it can not reach an object at motion planning time and so, the initial HATP plan will not take this into account (but when the action to pick the object not reachable failed, the robot will update its knowledge and so the new HATP plan). To access an object not reachable by it the robot can either navigate to another position (as in Fig. 7.7(c)), remove the object which blocks the access or get the object from the human (through handover).

The MF should allow the robot to learn in which case an object is really reachable by the robot and in which case another solution is preferred to get the object.

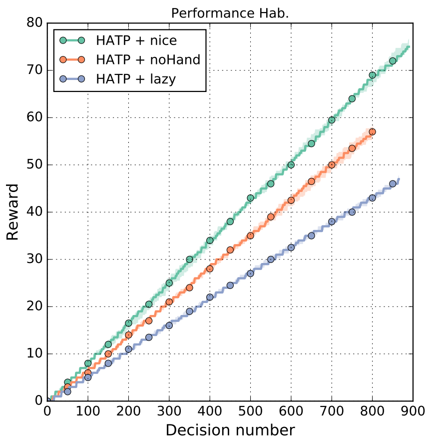
Different human behaviors Finally, in the previous task, one of the reason HATP was performing very well was that the human always executed the actions planned for him. In real life, even if the human is collaborative, he does not necessarily take the same decisions as the ones HATP took for him. In this task, we introduced three different kinds of human behavior:

— **The collaborative human:** he picks all objects accessible by him (with a priority for blue ones), throws the blue objects in the blue box and participates to all handover engaged by the robot.

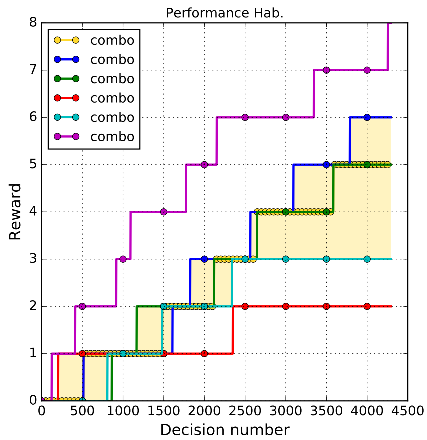
— **The anti-handover human:** he picks all the blue objects accessible by him, throws the blue objects in the blue box but does not participate to handover engaged by the robot (he does not react when the robot tries to execute a handover, the robot waits a few time and abort the action).

— **The lazy human:** he picks the blue objects accessible only by him (and not the ones the robot can access), throws the blue objects in the blue box and does not participate to handover engaged by the robot.

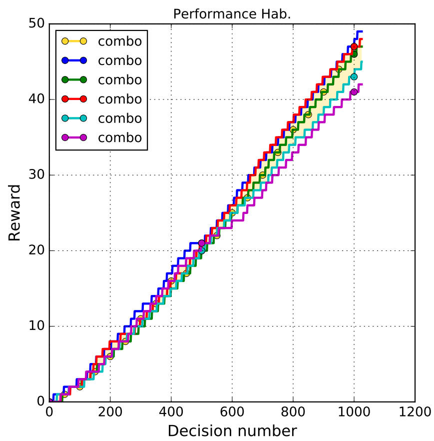
7.5.3 Results



(a) Mean cumulative reward for the system with only HATP. We can see that the system performs better with a collaborative human (nice), then with an human rejecting handover (noHand) and then with a lazy human (lazy).



(b) Mean cumulative reward for the system with only the MF. Different values for the MF parameters have been tested in order to find the best configuration in this task



(c) Mean cumulative reward for the system with the combination of both experts; Different values for the arbitration criteria parameters have been tested in order to find the best configuration in this task.

Figure 7.8: Mean cumulative reward for each conditions tested (HATP only, MF only and combination). The results are for 10 runs of approximatively 40 minutes in each conditions where the robot repeatedly fulfils the task.

We first tried the new architecture and task with HATP as the only expert. We can see in Fig.D.26(a) that, as expected, HATP performs better with a collaborative human which will have a behavior closer than the one its planned that with humans with less collaborative behaviors.

Then, we tested the system with the MF alone and with different parameters of the learning algorithm in order to get the better possible instantiation for this task. In a first step, we did it with a collaborative human and with only one possible initial set-up without geometrical complications. We can see in Fig. D.26(b) that, as expected, the MF alone performs poorly compared to HATP.

Then, we tested the combination of both experts. In a first step, we tested it with the collaborative human and with only one possible initial set-up without geometrical complications. We tested several parametrisations of the arbitration criteria in order to get the best implementation for this task. However, we noticed that, even when we put the system in the best possible situation, it performs barely as well as HATP in its worst case (the task was solved around 45 times in each cases, see Fig. D.26(c)). Indeed, with a more complex task, the bootstrap effect of HATP was not enough for the MF to learn a sufficiently good action policy. We tried with some runs way longer (several hours) but it was still not sufficient for the MF to learn a correct policy.

Moreover, even if we tried to reduce computation time by putting the meta controller upstream from the experts, the effect was not the one expected. Indeed, the meta controller here is probabilistic and so, even if the probability to choose the MF becomes higher than HATP, it can still happen for HATP to be chosen. In this case, a whole plan is computed by HATP even if we ask it only one action during the task. Consequently, the planning time remains the same than if HATP follows its plan alone to achieve the task.

7.6 Conclusion

In this chapter, we presented an architecture allowing to combine learning (a model free algorithm) and planning (a human-aware task planer HATP) during the robot decisional process. First results have shown that HATP allows to bootstrap the learning and so to quickly learn a consistent and acceptable behavior for the robot.

Secondly, we tried to show the benefits of the learning in the system. The result was not the one expected but we can still learn some lessons from this work and think of solutions to improve the system. One first possible modification would be to rework on the learning algorithm in order to study if there is method more adapted to this context. Then, another improvement would be to look for a new arbitration criteria between the two experts. Maybe a criteria with an hysteresis in order to reduce switches between experts in a task and allow them to have time to develop their own strategy (and not having one expert breaking the strategy the other tried to set-up) would be a good idea. Finally, one interesting idea would be to allow HATP to have a feedback on what is learned by the MF. Indeed, the knowledge of HATP concerning the actions is put off-line and is not updated during the interaction. For example, maybe the learning can provide the real time needed to execute an action or its probability of success given what was learned from previous interactions.

Despite the mixed results on the second system, we can still draw conclusions for this work:

— We have shown that the combination of HATP and the MF allows effectively to improve drastically the learning phase of the MF. Indeed, it allows the MF to quickly learn a policy and to adapt it to possible changes in the task.

— Thanks to its pertinent model of human-robot Joint Action, HATP allows the MF to learn whenever it is more efficient for the robot to *Wait* for an action of the human rather than trying to achieve the goal by itself.

— There are still some difficulties and this topic is far from being solved. However, the work presented here makes a step forward and gives clues on how to combine task learning and planning into future architecture.

Conclusion

Contributions

This manuscript presented several contributions to the field of human-robot cooperative task achievement. These contributions have been grouped in three parts:

- In a first part, we studied the basis of the Joint Action principles in order to build a supervisor for human-robot Joint Action.
  - In Chapter 1, we first studied the bibliography about Joint Action between humans in social science in order to identify what are the needed components for a robot decision adapted to Joint Action. Secondly, we looked how these components have already been applied in robotics and, finally, we studied how to articulate all these components and how, inspired by models from philosophy, we can build a robotics architecture for human-robot Joint Action.
  - In Chapter 2, we presented the supervisor for human-robot collaboration which has been developed and improved during the thesis. This supervisor is the main technical contribution of the thesis.
- In a second part, we focused on the achievement by a robot of a *Shared Plan* in collaboration with a human:
  - In Chapter 3, we presented how we endowed the robot with the ability to estimate the humans mental states, not only about the environment, but also concerning the state of the task and more particularly of the Shared Plan. We have also shown how we used these mental states to allow the robot to better communicate about divergent beliefs during Shared Plan execution.
  - In Chapter 4, we described the work done in order to allow the robot to gain acceptability and fluency during Shared Plan achievement by working with more flexible Shared Plans. We first identified the needed decisions during Shared Plan elaboration and execution and we endowed the robot with the ability to decide which decisions should be taken at planning time and which ones are better postponed to execution time. Then, we allowed the robot to take these decision by smoothly adapting to the human choices.
  - Finally, in Chapter 5, we evaluated the new implemented system for the achievement of human-robot Shared Plan. This evaluation has been done quantitatively in simulation but also qualitatively with a user study with the real robot. Both evaluations highlighted the pertinence of the improvements brought to the system. A questionnaire allowing to evaluate the users feelings about a collaboration with a robot has been developed and validated (in terms of intern coherence) in the context of the

user study. Moreover, this user study allows us to get more insights on experiments with naive subjects.

In the last part, we presented other contributions to the domain:

— In Chapter 6, we studied the non-verbal behavior, and more especially the head -gaze- behavior, needed during Joint Action between humans and between a human and a robot. We identified the needed components of a robot head behavior adapted to the Joint Action and investigated more deeply some of them with an on-line video based study. Finally, we presented how these components can be implemented into a robot head behavior architecture.

— In Chapter 7, inspired from studies in neuroscience, we combined learning and planning for high level decisions during human-robot Joint Action. The idea being to take advantage from both techniques in order to come up with decision level which is able to quickly learn how to smoothly adapt to the human choices during Joint Action execution.

Future works and improvements

All these contributions made a step toward the aim of a full autonomous robot able to work jointly with humans in the context of Joint Action. However there is still plenty room for improvements and plenty other enhancements to bring to attain this goal. Concerning the work presented in this manuscript:

— Concerning the work on Shared Plan achievement, there is still plenty of possible modifications in order for the execution to be even more flexible and adapted to the human choices. We can mention, among others, the possibility to deal with the temporary absence of the human by computing plans where the robot tries to achieve the maximum on its own for a time while keeping in mind that the human should return in order to achieve the action only him can perform (instead of failing because nobody can achieve these actions anymore).

— Concerning the work on the robot head behavior, the next step is to implement the proposed architecture on the robot. Once this done, it will be interesting to evaluate this architecture through a user study. It will also be interesting to make the link between the signals transmitted and the estimated mental states of the humans. Indeed, when the robot transmits a signal, in addition to make sure that the signal has been well received, it should update the mental state of its partner with the information that he received the signal and all what it implies. Finally, this work should be extended to the rest of the robot non-verbal behavior.

— Concerning the work on the combination of learning and planning, there is still work to do to have a conclusive architecture for high level decisions. Moreover, it can be interesting to improve the combination by allowing the planner to learn costs and execution times and maybe probabilities of success as well from the learning in order to update its model and come up with more appropriate plans.

We can also mention, among others, several other improvements to bring to the current system:

— it could be interesting focus on issues relative to engagement during human-robot Joint Action. The robot should be able, in addition to detect to the engagement of its human partners on the current task, to deal with small distraction of its partner (without thinking the partner is disengaged from the goal) or interruptions of the task to perform by another one (imagine you give a long task to perform to your robot in collaboration with your son, you want to be able to interrupt it the time it brings you a beer).

— It can also be interesting to link humans actions perception to the Shared Plan execution. Indeed, knowing what are the possible actions of the human and what are the actions we expect him to do can help the robot to have a more robust detection of humans actions.

— Finally, it could also be interesting to work on the link between dialogue and supervision during Joint Action. There are two possibilities, and in my point of view needed approaches to do it. First the dialogue can be seen as a tool to the Joint Action execution (e.g. as started in this thesis by giving appropriate information at the right time). Secondly, the dialogue can also be seen as a full Joint Action, with the supervision helping to perform actions supporting the dialogue (e.g. pointing a point of interest). Moreover, a strong coordination should be done between the non-verbal behavior needed during dialogue and the one needed during Shared Plan execution.

Appendix A

Terms of the formalization

TS**:** state of the task from the robot point of view.

$$TS =< g_R, SP, WS >$$

$MS(H)$**:** mental state of the human H.

$$MS(H) =< g_H, g_R(H), SP(H), WS(H) >$$

Goals:
g_R**:** current goal of the robot
g_H**:** current goal of the human H
$g_R(H)$**:** current goal of the robot from the human H point of view

$$g =< Name_g, Actors_g, Params_g, Obj_g >$$

$Name_g$**:** identifier of the goal g
$Actors_g$**:** actors of the goal g
$Params_g$**:** parameters of the goal g
Obj_g**:** objectives of the goal g
$label_g$**:** state of a goal g already over (DONE or ABORTED)

Shared Plan:
SP**:** current Shared Plan
$SP(H)$**:** current Shared Plan from the human H point of view

$$SP =< id_p, A_p, L_p >$$

id_p**:** identifier of the plan p
A_p**:** actions of the plan p

$$A_p =< A_{prev}, A_{cur}, A_{next}, A_{later} >$$

L_p**:** links between actions for the plan p

$$l \in L_p = \langle prev_l,\ next_l \rangle$$

$prev_l$**:** identifier of the action to execute first
$next_l$**:** identifier of the action to execute next (after $prev_l$)

Actions sets:

A_{prev}**:** actions already executed

$A_{prev}(H)$**:** actions already executed from the human H point of view

A^R_{prev}**:** actions already executed by the robot

$A^R_{prev}(H)$**:** actions already executed by the robot from the human H point of view

A^H_{prev}**:** actions already executed by the human H

$A^H_{prev}(H)$**:** actions already executed by the human H from the human H point of view

A_{cur}**:** actions in progress

$A_{cur}(H)$**:** actions in progress from the human H point of view

A^R_{cur}**:** actions in progress and executed by the robot

$A^R_{cur}(H)$**:** actions in progress and executed by the robot from the human H point of view

A^H_{cur}**:** actions in progress and executed by the human H

$A^H_{cur}(H)$**:** actions in progress and executed by the human H from the human H point of view

A_{next}**:** actions from the plan which need to be performed

$A_{next}(H)$**:** actions from the plan which need to be performed from the human H point of view

A^R_{next}**:** actions from the plan which need to be performed by the robot

$A^R_{next}(H)$**:** actions from the plan which need to be performed by the robot from the human H point of view

A^H_{next}**:** actions from the plan which need to be performed by the human H

$A^H_{next}(H)$**:** actions from the plan which need to be performed by the human H from the human H point of view

A^X_{next}**:** actions from the plan which need to be performed and are not yet allocated

$A^X_{next}(H)$**:** actions from the plan which need to be performed and are not yet allocated from the human H point of view

A_{later}**:** actions from the plan which need to be performed later

$A_{later}(H)$**:** actions from the plan which need to be performed later from the human H point of view

A^R_{later}**:** actions from the plan which need to be performed later by the robot

$A^R_{later}(H)$**:** actions from the plan which need to be performed later by the robot from the human H point of view

A^H_{later}**:** actions from the plan which need to be performed later by the human H

$A^H_{later}(H)$**:** actions from the plan which need to be performed later by the human H from the human H point of view

A^X_{later}**:** actions from the plan which need to be performed later and are not yet allocated

$A^X_{later}(H)$**:** actions from the plan which need to be performed later and are not yet allocated from the human H point of view

Action:

$$a =< id_a,\ Name_a,\ Ag_a,\ Params_a,\ Precs_a,\ Effects_a >$$

id_a**:** identifier of the action a
$Name_a$**:** name of the action a
Ag_a**:** actors of the action a
$Params_a$**:** parameters of the action a
$Precs_a$**:** preconditions of the action a
$Effects_a$**:** effects of the action a
$label_a$**:** state of an action a already executed (DONE, FAILED or ABORTED)

World State:
WS**:** current world state from the robot point of view
$WS(H)$**:** current world state from the human H point of view

$$p \in WS =< entity, attribute, value >$$

Appendix B

Questionnaire of the on-line video based study for the robot head behavior

B.1 Anticipation of robot actions

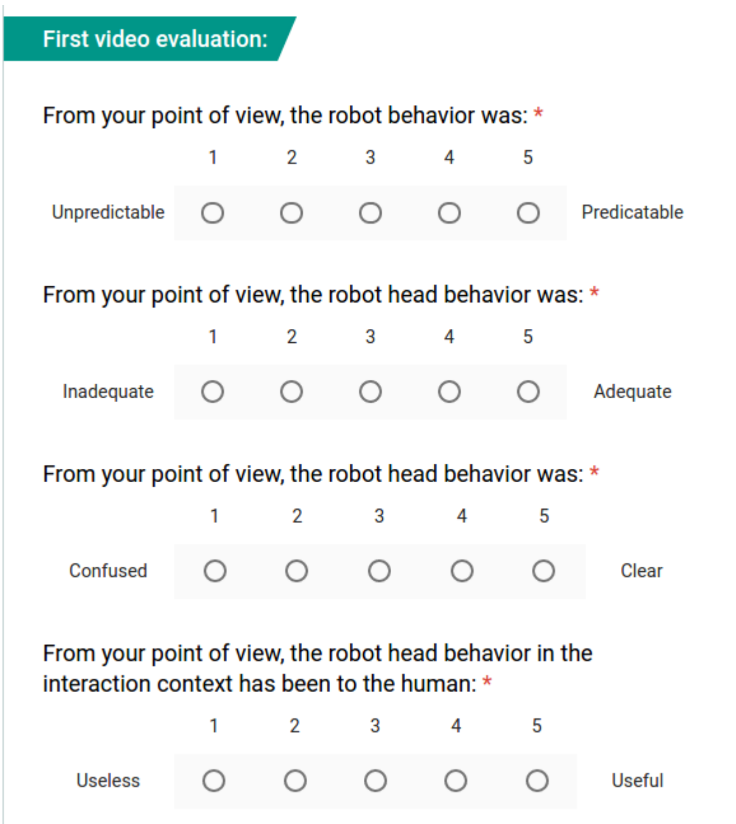
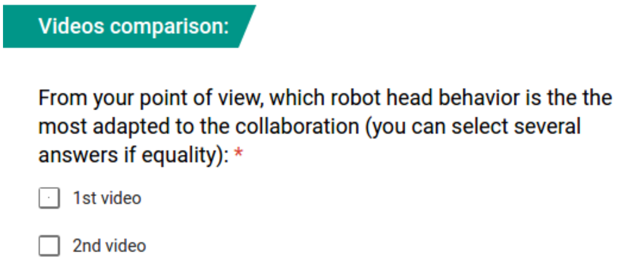


Figure B.1: Questions asked to participant for each videos of the scenario concerning the anticipation of robot actions.



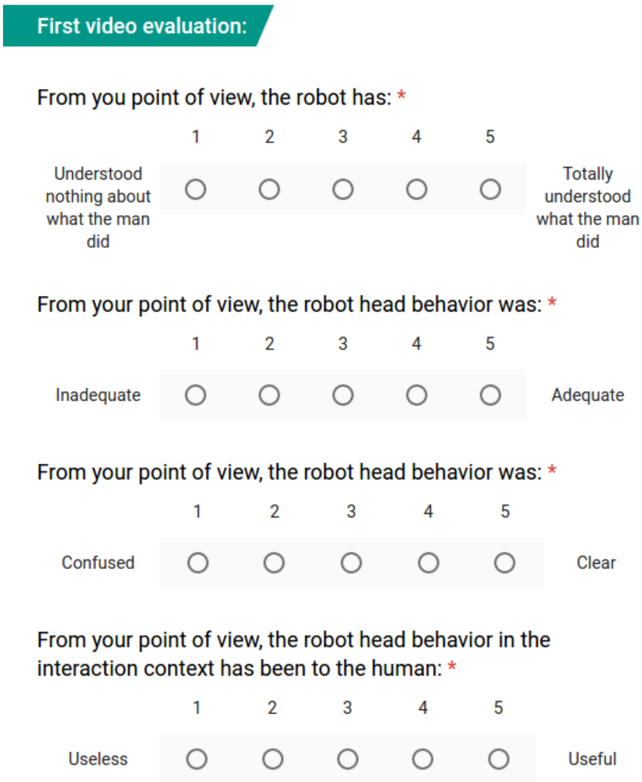
Videos comparison:

From your point of view, which robot head behavior is the the most adapted to the collaboration (you can select several answers if equality): *

- [ ] 1st video
- [ ] 2nd video

Figure B.2: Question asked to participant to compare the two videos they watched of the scenario concerning the anticipation of robot actions.

B.2 Tracking human's activity



First video evaluation:

From you point of view, the robot has: *

	1	2	3	4	5	
Understood nothing about what the man did	○	○	○	○	○	Totally understood what the man did

From your point of view, the robot head behavior was: *

	1	2	3	4	5	
Inadequate	○	○	○	○	○	Adequate

From your point of view, the robot head behavior was: *

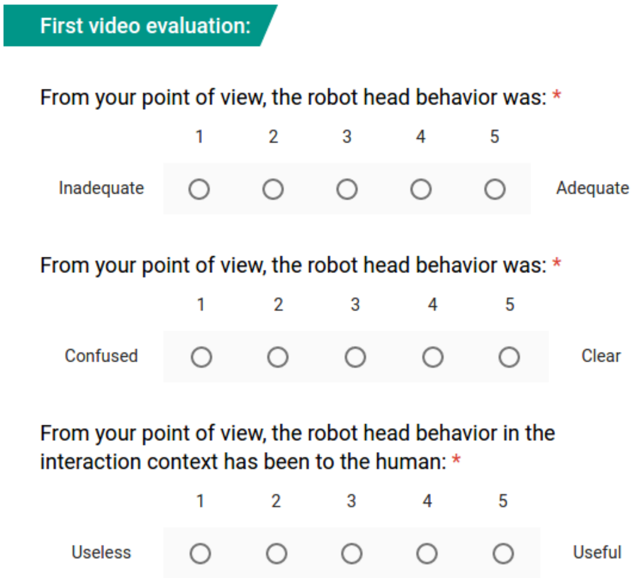
	1	2	3	4	5	
Confused	○	○	○	○	○	Clear

From your point of view, the robot head behavior in the interaction context has been to the human: *

	1	2	3	4	5	
Useless	○	○	○	○	○	Useful

Figure B.3: Questions asked to participant for each videos of the scenarios concerning the tracking of human's activity.

B.3 Helping the human to perform his next action


First video evaluation:

From your point of view, the robot head behavior was: *

	1	2	3	4	5	
Inadequate	○	○	○	○	○	Adequate

From your point of view, the robot head behavior was: *

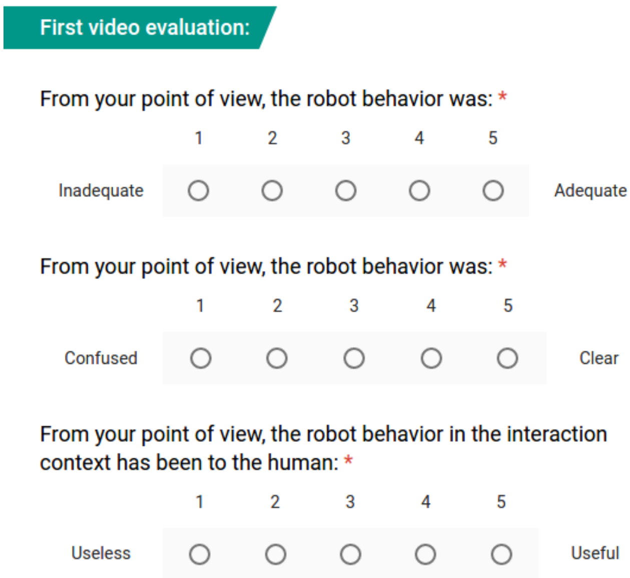
	1	2	3	4	5	
Confused	○	○	○	○	○	Clear

From your point of view, the robot head behavior in the interaction context has been to the human: *

	1	2	3	4	5	
Useless	○	○	○	○	○	Useful

Figure B.4: Questions asked to participant for each videos of the scenarios concerning the help of the human to perform his next action.

B.4 Finding the priority target


First video evaluation:

From your point of view, the robot behavior was: *

	1	2	3	4	5	
Inadequate	○	○	○	○	○	Adequate

From your point of view, the robot behavior was: *

	1	2	3	4	5	
Confused	○	○	○	○	○	Clear

From your point of view, the robot behavior in the interaction context has been to the human: *

	1	2	3	4	5	
Useless	○	○	○	○	○	Useful

Figure B.5: Questions asked to participant for each videos of the scenarios concerning the signalling of human's actions.

Appendix C

Questionnaires of the user-study

C.1 Reminder questionnaire

Some questions...

Did the robot asked you questions?
◯ Oui ◯ Non

Did the robot gave you information?
◯ Oui ◯ Non

Did the robot imposed you its decisions?
◯ Oui ◯ Non

Are you sure that all the objects have been scanned?
◯ Oui ◯ Non

Did the robot participate to store the red objects?
◯ Oui ◯ Non

Suggestions :
____
____
____

Figure C.1:

C.2 General questionnaire

Participant number :
Condition :

The aim of this questionnaire is to evaluate the different behaviors of the robot betwenn the four trials. Please insert a check in the box which corresponds the most to your feeling, do not hesitate to help yourself whith the questionnaire « some questions ».

In your opinion, the robot's verbal interventions was :

Incomprehensible — Clear

Insufficient — Sufficient

Superfluous — Pertinent

Other: ………………………………

In your opinion, the robot's acts was :

Unsuitable — Appropriate

Useless — Usefull

Unpredictable — Predictable

Other : ………………………………

In your opinion, the collaboration with the robot to perform the task was :

Restrictive — Adaptative

Unnecessary — Necessary

Unsettling — Satisfactory

Annoying — Acceptable

Insecure — Secure

Other: ………………………………

Figure C.2:

In your opinion, as a rule, the interaction was:

Negative	Positive
Simple	Complicated
Not practical	Practical
Unpredictable	Predictable
Ambiguous	Clear

Other : ………………………………

In your opinion, the robot is rather :

Machinelike	Humanlike
Artificial	Lifelike
Stagnant	Lively
Apathetic	Responsive
Unpleasant	Pleasant
Annoying	Lovely
Stupid	Intelligent
Incompetent	Competent

Other : ………………………………..

As a result of this interaction, circle the icon you identify best with:

Figure C.3:

Appendix D

French extended abstract

D.1 Introduction

Dans les années 40, des chercheurs inventent les premières machines appelées ordinateurs. En 1956, à la conférence de Darmouth, le domaine de l'intelligence artificielle est reconnu et les premiers robots arrivent rapidement dans notre environnement. Certains de ces robots vont devoir évoluer avec les Hommes ou dans leur entourage. Entre autres, les robots "co-workers" en industrie ou les robots sociaux [Dautenhahn 2007]. Le but de cette thèse est de se rapprocher de robots qui peuvent agir conjointement avec les Hommes de manière naturelle, fluide et efficace. On se concentre ici sur les problématiques liées aux processus décisionnels durant l'action conjointe Homme-Robot.

Dans un premier temps, basé sur une étude bibliographique des éléments nécessaires à l'action conjointe entre Hommes ainsi que sur des travaux existants en interaction Homme-Robot, les différents éléments nécessaires à l'action conjointe Homme-Robot seront identifiés ainsi que la manière dont ils peuvent s'articuler dans une architecture. Puis, l'architecture du superviseur, contribution technique principale de la thèse, sera présentée. Dans un second temps, mes travaux concernant l'amélioration de la gestion des plans partagés par le robot durant l'action conjointe seront présentés. La première amélioration concerne la prise en compte des états mentaux des Hommes durant l'exécution de plans partagés. La seconde contribution concerne le report de certaines décisions prises initialement par le robot durant l'élaboration du plan et à l'exécution afin d'obtenir une gestion plus flexible des plans partagés. L'évaluation de ces deux contributions en simulation et à l'aide d'une étude utilisateur sera également présentée. Finalement, dans un troisième temps, deux autres contributions à l'action conjointe Homme-Robot seront présentées. La première concerne la gestion du comportement non-verbal et plus précisément de la tête du robot. La seconde concerne l'association d'un système d'apprentissage à un système de planification dans le cadre de la prise de décision haut niveau.

D.2 De l'action conjointe entre Hommes à la supervision pour l'interaction Homme-Robot

D.2.1 De l'action conjointe entre Hommes à l'action conjointe Homme-Robot

D.2.1.1 Théorie de l'action conjointe

L'action conjointe a été décrite par [Sebanz 2006] comme :

> *n'importe quelle forme d'interaction sociale où deux individus ou plus coordonnent leurs actions dans l'espace et le temps pour apporter un changement dans l'environnement.*

Plusieurs prérequis sont nécessaires pour que deux individus réalisent avec succès une action conjointe.

La première chose requise est que ces individus partagent un but et l'intention d'achever ce but. [Tomasello 2005] défini un but comme la représentation d'un état désiré par un agent et une intention comme un plan d'action qu'un agent s'engage à exécuter pour réaliser le but (basé sur le travail de Bratman [Bratman 1989]). Dans le cas de l'action conjointe, une des définitions les plus reconnues est celle de Bratman [Bratman 1993] qui présente trois conditions pour que deux individus partagent une *intention jointe* d'accomplir un but G :

1. Chaque individu a l'intention d'accomplir G.
2. Chaque individu a cette intention en accord avec 1 et les parties du plan partagé de 1.
3. 1 et 2 sont une connaissance commune entre chaque individu.

Cette définition est reprise et illustrée par Tomasello et al. dans [Tomasello 2005] (Fig. D.1). Ils définissent un *but partagé* comme une représentation d'un état désiré plus la connaissance que le but va être réalisé en collaboration et une *intention jointe* comme un plan partagé auquel les agents se sont engagés pour réaliser le but. Concernant ce plan partagé, cette notion a été introduite et formalisée par Grosz and Sidner [Grosz 1988]. Leur définition suggère que chaque agent ne connaît pas nécessairement le plan entier mais seulement la partie qui le concerne et les parties en intersection avec celles de ses partenaires.

Un deuxième prérequis de l'action jointe est que chaque agent doit être capable de percevoir et de prédire les actions de ses partenaires et leurs effets. A partir des travaux de [Sebanz 2006], [Pacherie 2011] et [Obhi 2011] nous avons identifié plusieurs capacités nécessaires à ces prédictions :

— **L'attention jointe :** la capacité d'un agent à diriger son attention vers le même objet que ses partenaires de manière à partager la même représentation de l'environnement et des événements.

— **Observation de l'action :** plusieurs études ont montré que quand quelqu'un observe une autre personne réaliser une action, une représentation de

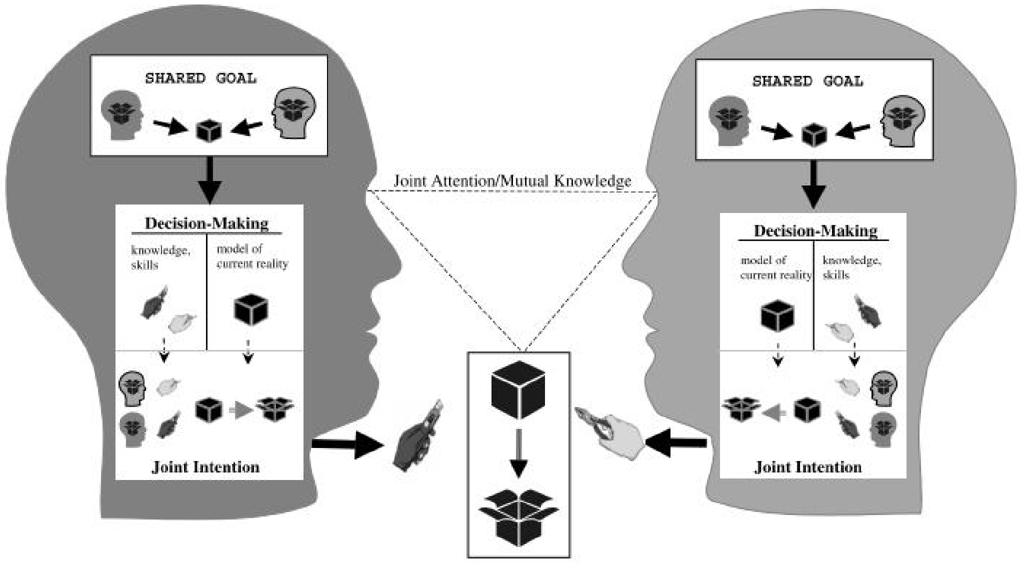


FIGURE D.1 – Exemple d'une activité collaborative par Tomasello et al. Ici les deux hommes ont pour **but partagé** d'ouvrir la boite ensemble. Ils ont choisi un moyen d'atteindre ce but qui prend en compte les capacités de chaque agent et ont donc une *intention jointe.*

cette action est formée par l'observateur par ce qu'on appelle les *neurones miroirs* et permet de prédire les effets de l'action [Rizzolatti 2004].
- **Co-représentation :** avoir une représentation de son partenaire (son but, ses capacités, ses connaissances, etc...) permet de prédire ses actions futures.
- **Agency :** la capacité d'attribuer les effets d'une action au bon acteur.

Grâce à ces capacités, plusieurs prédictions peuvent être effectuées :
- **Quoi :** prédire quelle action un agent va réaliser.
- **Quand :** prédire quand une action va avoir lieu et combien de temps elle va durer pour mieux se coordonner dans le temps.
- **Où :** prédire les futures positions de ses partenaires pour mieux se coordonner dans l'espace.

Finalement, pendant l'action conjointe, les agents doivent être capables de coordonner leurs actions dans le temps et l'espace. Deux sortes de coordination sont définies dans [Knoblich 2011] :
- **La coordination émergente :** qui a lieu intentionnellement. La coordination émergente peut être due à plusieurs mécanismes tels que *l'entraînement* [Richardson 2007], des *affordances* communes [Gibson 1977] ou la perception d'une action.
- **La coordination planifiée :** qui est, elle, intentionnelle. Pour cela, les agents peuvent modifier leur comportement avec ce qui est défini par [Vesper 2010] comme des *coordination smoothers* (mouvements plus prédictibles, signaux de coordination, etc...) ou utiliser la communication verbale

ou non-verbale [Clark 1996].

D.2.1.2 Comment doter un robot des capacités nécessaires à l'action conjointe ?

Plusieurs travaux ont déjà été réalisés afin de doter le robot des capacités nécessaires à la réalisation d'une action conjointe avec l'Homme.

Dans un premier temps, le robot doit être capable de s'engager dans une action conjointe, de choisir un but. Ce but peut être imposé par l'utilisateur mais le robot doit aussi être capable de pro-activement proposer son aide. Pour faire cela, plusieurs travaux ont été réalisés concernant la reconnaissance de plans [Ramırez 2009, Bui 2003, Singla 2011] et d'intentions [Breazeal 2009, Baker 2014]. Le robot doit également être capable de choisir entre différents buts possibles. Pour faire cela le domaine de l'intelligence artificielle a commencé à proposer plusieurs solutions [Ghallab 1994, Lemai 2004, Roberts 2016]. Une fois le robot engagé dans une action conjointe, il doit être capable de surveiller l'engagement de ses partenaires. Des réponses à ce problème ont été données en utilisant les signaux visuels et gestes [Sanghvi 2011] ou le contexte et les états mentaux [Salam 2015]. Enfin, une fois le robot engagé dans un but, il doit être capable d'obtenir un plan partagé. Ce plan peut être imposé par l'utilisateur et le robot doit alors être capable de le comprendre [Pointeau 2014, Mohseni-Kabir 2015] et éventuellement de le retransmettre [Petit 2013, Sorce 2015]. Le plan peut aussi être construit en collaboration [Allen 2002] ou élaboré par le robot [Cirillo 2010, Lallement 2014]. Si le robot élabore le plan, il doit également être capable de communiquer à son sujet [Milliez 2016b].

Afin de mieux communiquer et travailler avec l'Homme, le robot doit être capable d'aligner sa représentation du monde (données en x, y, z venant des capteurs) avec celle de l'Homme (relations sémantiques entre objets). Ce processus a été étudié et s'appelle *l'ancrage* [Coradeschi 2003, Mavridis 2005, Lemaignan 2012]. Le robot doit également être capable de représenter son environnement non seulement de son point de vue mais aussi de celui de ses partenaires. La prise de perspective du robot [Breazeal 2006, Milliez 2014] peut être utilisée pour résoudre des situations ambiguës [Ros 2010a], mieux interagir durant le dialogue [Ferreira 2015] ou reconnaître et interpréter les actions de l'Homme [Baker 2014, Nagai 2015].

Finalement, le robot doit être capable de se coordonner avec l'Homme. A un haut niveau, l'Homme et le robot doivent coordonner leurs actions afin de réaliser le plan partagé avec succès. Plusieurs systèmes permettent de faire cela tels que *Chaski* [Shah 2011], *Pike* [Karpas 2015] ou *SHARY* [Clodic 2009]. Pour faire cela, le robot doit se reposer sur ses capacités de communication verbale [Roy 2000, Lucignano 2013, Ferreira 2015] et non-verbale [Breazeal 2005, Boucher 2010, Mutlu 2009b, Hart 2014]. A un plus bas niveau, le robot doit se coordonner avec l'Homme durant l'exécution d'actions telles que le transfert d'un objet. Cela représente plusieurs challenges tels que trouver des postures acceptables par l'Homme [Cakmak 2011, Dehais 2011, Mainprice 2012], approcher un Homme [Walters 2007] ou produire des trajectoires lisibles et prédictibles [Sisbot 2012, Kruse 2013].

D.2.1.3 Une architecture trois niveaux

Nous avons vu précédemment les prérequis pour l'action conjointe entre Hommes et Homme-Robot. Nous allons maintenant voir comment ces éléments se combinent en une architecture trois niveaux.

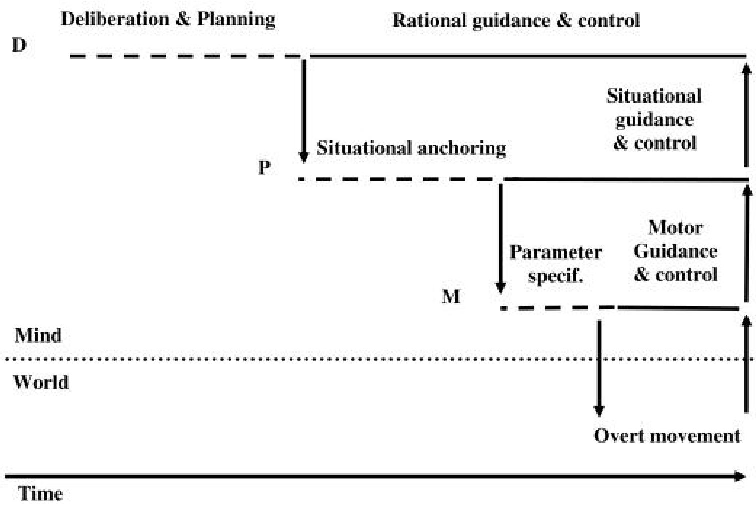


FIGURE D.2 – Les niveaux de Pacherie mis en cascade. Chaque niveau contrôle l'action à un niveau différent.

En ce qui concerne l'action conjointe entre Hommes, Pacherie [Pacherie 2011] défend le fait que les processus liés à l'action conjointe se décomposent en trois niveaux qui ont chacun leur rôle (Fig. D.2) :

- **Shared Distal Intention :** c'est le niveau le plus haut. Ce niveau est responsable de la formation d'une *intention jointe* et de la gestion du plan partagé.
- **Shared Proximal Intention :** ce niveau a la responsabilité d'ancrer les actions reçues du niveau supérieur dans le contexte actuel. Cela doit être fait de manière coordonnée avec les partenaires de l'action conjointe.
- **Coupled Motor Intention :** c'est le niveau le plus bas. Il est responsable des commandes moteurs des agents. Il s'occupe de la coordination spatio-temporelle au niveau le plus précis.

10 années avant que Pacherie développe ses idées concernant l'architecture trois niveaux, le domaine de la robotique autonome concevait intuitivement des architectures avec trois niveaux très similaires comme dans [Alami 1998] où l'on retrouve les niveaux :

- **Niveau décisionnel :** il est responsable de la production et la supervision du plan d'action. Il peut être comparé au niveau *Distal* de Pacherie.

— **Niveau exécutionnel :** il a la responsabilité de choisir, paramétrer et synchroniser les différentes fonctions nécessaires à l'exécution des actions venant du niveau décisionnel. Il peut être comparé au niveau *Proximal* de Pacherie.
— **Niveau fonctionnel :** il comprend toutes les foncions bas niveau d'action et de perception du robot. Il peut être comparé au niveau *Motor* de Pacherie.

Cette architecture a été développée et adaptée au domaine de l'interaction Homme-robot. Récemment, nous avons présenté dans [Devin 2016b] une version théorique d'une architecture adaptée à l'action conjointe Homme-robot (Fig. D.3).

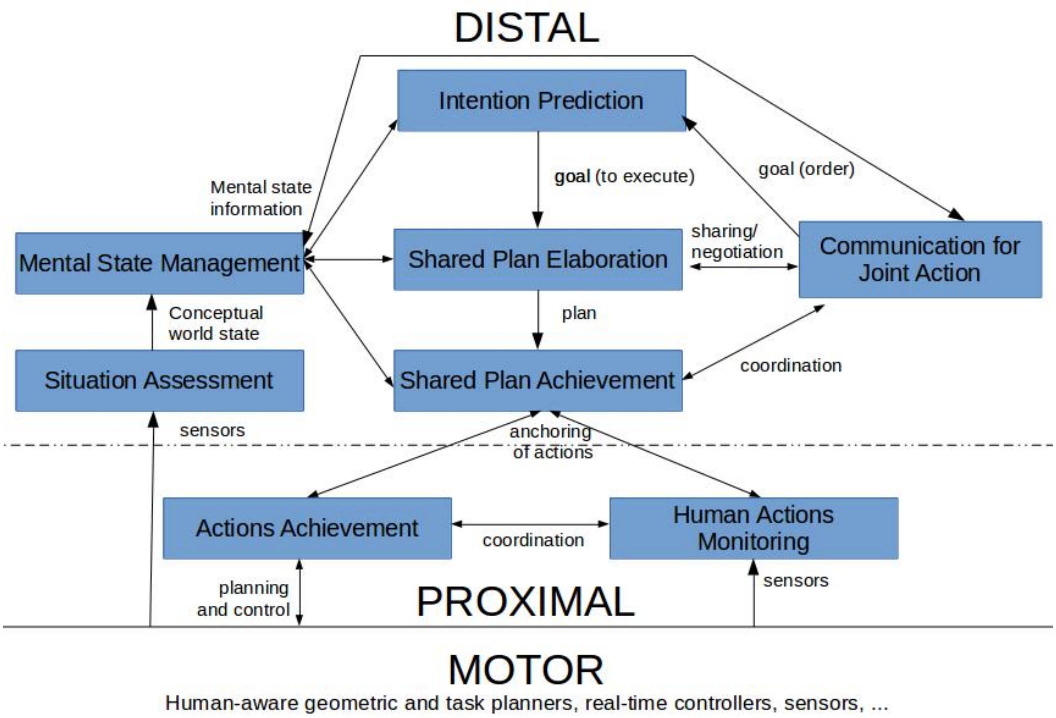


FIGURE D.3 – Architecture récente pour l'action conjointe Homme-robot. L'architecture est organisée autour des trois niveaux définis par Pacherie.

D.2.2 Supervision pour l'interaction Homme-Robot

D.2.2.1 Rôle du superviseur dans l'architecture globale

Le superviseur faisant l'objet de cette thèse fait partie d'une architecture globale pour l'interaction Homme-Robot développée au LAAS-CNRS. Une version simplifiée de cette architecture peut être trouvée Fig. D.4.

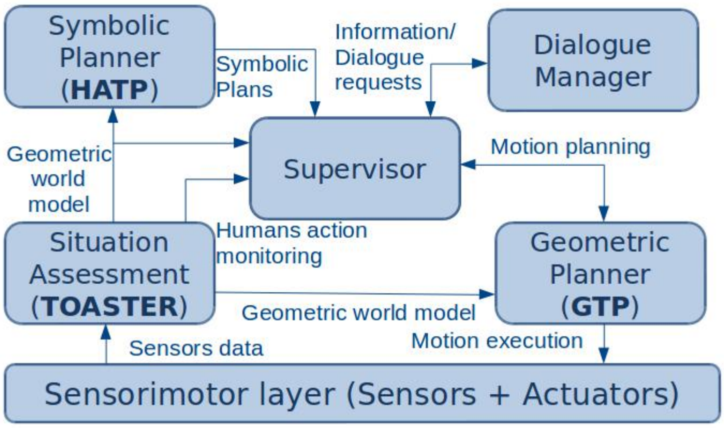


FIGURE D.4 – Architecture globale pour l'interaction Homme-robot développée au LAAS-CNRS.

Cette architecture est composée de :

— **Un niveau sensorimoteur :** qui contient les modules bas niveau du robot lui permettant de gérer ses capteurs et actionneurs.

— **TOASTER :** un module permettant au robot de représenter et maintenir un état du monde symbolique de son point de vue ainsi que de celui de ses partenaires.

— **GTP :** un planificateur géométrique permettant au robot d'effectuer des actions en prenant en compte le confort et la sécurité de l'Homme.

— **HATP :** un planificateur symbolique permettant au robot de calculer des plans pour lui même et pour ses partenaires afin d'atteindre un but donné.

— **Un module de dialogue :** permettant au robot de communiquer avec l'Homme.

— **Un superviseur :** ayant la charge de superviser l'activité du robot en coordonnant les autres modules. Il choisit le but du robot, veille au bon déroulement du plan partagé, choisit quand exécuter une action et comment communiquer.

L'architecture interne du superviseur peut être trouvée Fig. D.5. Il est composé de plusieurs modules :

— **Un gestionnaire de but :** permettant au robot de choisir quel but exécuter à chaque moment.

— **Un élaborateur de plan :** permettant au superviseur de communiquer avec HATP afin d'obtenir un plan partagé.

— **Un mainteneur de plan :** permettant au robot de suivre l'évolution du plan partagé.
— **Un estimateur d'états mentaux :** permettant au robot d'estimer les états mentaux de ses partenaires humains concernant le plan partagé.
— **Un module de décision :** permettant au robot de choisir quand exécuter une action ou donner une information.
— **Un module d'exécution d'action :** permettant au superviseur de surveiller la bonne exécution des actions par le robot.
— **Un module de communication non-verbale :** permettant de gérer pour le moment uniquement la tête du robot.

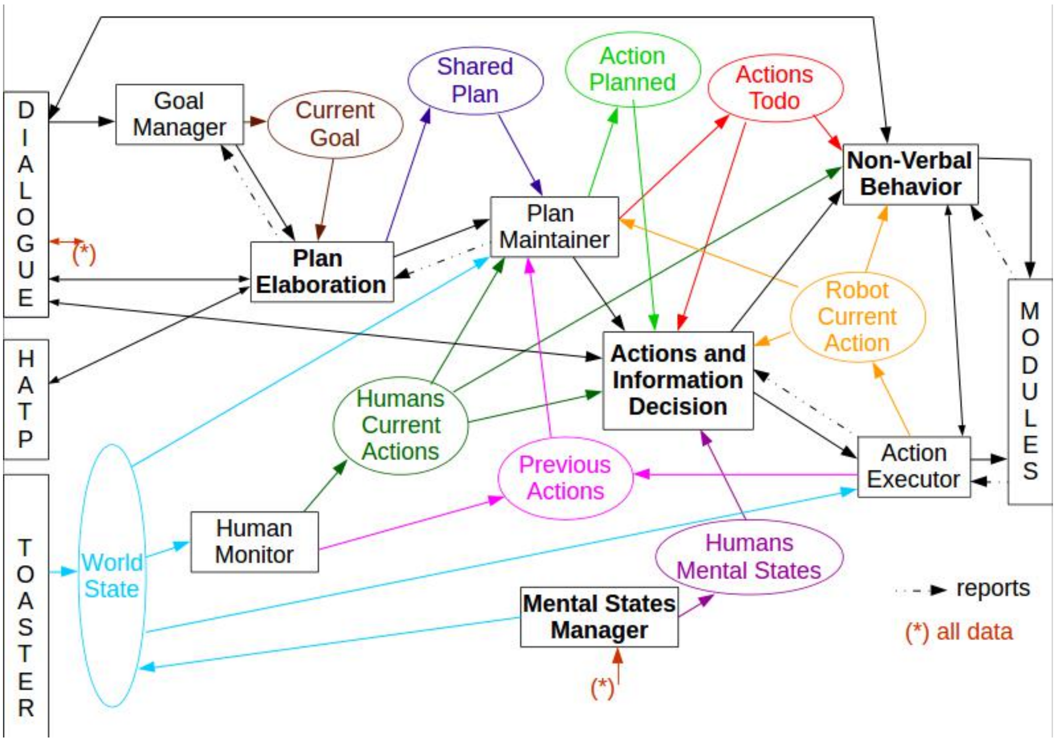


FIGURE D.5 – Architecture interne du superviseur. Les modules en gras sont traités dans ce manuscrit.

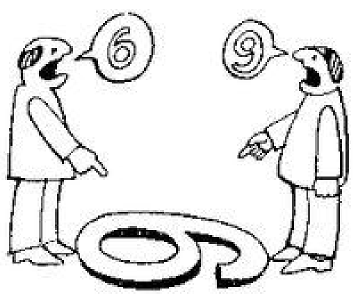
D.3 Les plans partagés durant l'action conjointe Homme-Robot

D.3.1 Prendre en compte les états mentaux pendant l'exécution de plans partagés

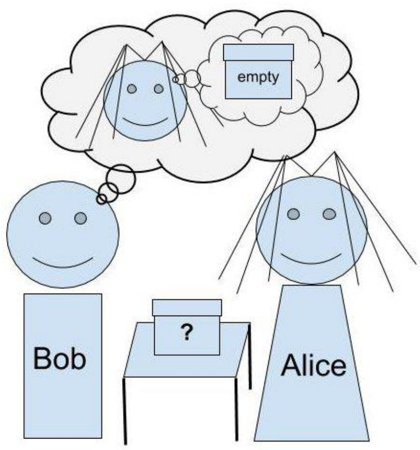
D.3.1.1 Motivations et précédents travaux

Quand le robot interagit avec un Homme, il est important qu'il ne le considère pas comme un outil ou un obstacle mais qu'il prenne en compte ses sentiments et son confort et donc son point de vue notamment lors de l'exécution d'un plan partagé.

La théorie de l'esprit désigne la capacité qu'ont les humains de reconnaître et s'attribuer des états mentaux en comprenant que les autres personnes peuvent avoir des connaissances et sentiments différents des leurs et de prendre en compte ces états mentaux pendant la prise de décision. La théorie de l'esprit a beaucoup été étudiée dans les sciences sociales [Baron-Cohen 1985, Premack 1978], notamment la notion de prise de perspective qui désigne la capacité d'une personne à prendre le point du vue d'une autre personne [Tversky 1999, Flavell 1992]. Deux niveaux de prise de perspective sont définis dans [Flavell 1977]. La prise de perspective perceptuelle désigne la capacité d'une personne à comprendre que les autres ont une représentation du monde différente de la sienne (fig D.6(b)). La prise de perspective conceptuelle désigne la capacité d'une personne à attribuer des croyances et connaissances à une autre personne (fig D.6(b)).



(a) Prise de perspective perceptuelle : deux individus peuvent avoir deux représentations différentes de leur environnement.



(b) Prise de perspective conceptuelle : ici Bob attribue à Alice une connaissance concernant la boite : il pense qu'Alice pense que la boite est vide.

FIGURE D.6 – Illustration de la prise de perspective perceptuelle et conceptuelle.

En robotique, plusieurs travaux ont pour but de doter le robot de capacités liées à la théorie de l'esprit. Un des premiers travaux sur ce sujet est celui de Scassellati où il propose un modèle pour adapter deux modèles des sciences sociales [Leslie 1984, Baron-Cohen 1997] afin d'implémenter la théorie de l'esprit en robotique [Scassellati 2002]. Plusieurs travaux ont permis aux robots de se doter de capacités de prise de perspective [Berlin 2006, Hiatt 2010, Milliez 2014]. Ces capacités ont été utilisées dans plusieurs travaux visant par exemple à mieux reconnaître et comprendre les actions de l'Homme [Johnson 2005a, Baker 2014, Nagai 2015] ou pour résoudre des situations ambiguës [Breazeal 2006]. Des travaux ont été réalisés pour prendre en compte le point de vue de l'homme durant l'élaboration d'un plan partagé [Warnier 2012], cependant, aucun ne concerne l'exécution de ce plan. Cette partie de la thèse a pour but de commencer à combler ce manque.

D.3.1.2 Estimation des états mentaux

Dans un premier temps le robot doit être capable d'étendre l'estimation des états mentaux de ses partenaires (qui concernait précédemment les connaissances sur l'environnement) aux connaissances concernant la tâche en cours et le plan partagé. Les algorithmes développés permettent au robot d'estimer les états mentaux de l'Homme concernant :

— **l'état du monde :** en plus de l'estimation des connaissances de l'Homme concernant l'état du monde observable venant de la prise de perspective (e.g. un objet est sur un autre objet) le robot est capable d'estimer les connaissances de l'Homme concernant l'état du monde non-observable (e.g. une boite est vide ou remplie) en se basant principalement sur les effets des actions.

— **le plan partagé :** en se basant sur ce que l'Homme peut observer, le robot est capable d'estimer ses connaissances concernant les actions en cours ou passées. Grâce à cette estimation et à ses propres connaissances concernant le plan partagé, le robot est capable d'estimer les connaissances de l'Homme concernant l'état du plan (e.g. quelles actions doivent être exécutées).

— **le but :** le robot est capable d'estimer des connaissances basiques de l'Homme concernant l'état du but en cours (e.g. si il est achevé) en se basant sur ses connaissances sur l'état du monde.

D.3.1.3 Utilisation des états mentaux durant l'exécution du plan partagé

Une fois les états mentaux de ses partenaires estimés, le robot doit être capable de correctement les utiliser afin de communiquer durant l'exécution du plan partagé quand une différence apparaît entre les connaissances du robot et celles de l'Homme. En effet, le robot doit fournir à l'Homme les informations dont il a besoin pour réaliser la tâche sans pour autant être trop verbeux en informant l'Homme à propos de tout et n'importe quoi. Pour cela, nous avons développé plusieurs comportements pour le robot :

— en accord avec la notion de *weak achievement goal* de [Cohen 1991], si le robot détecte une différence entre ses connaissances concernant l'état du but en cours et celle d'un de ses partenaires, le robot va informer ce partenaire à propos de cette différence.

— lorsque le robot estime qu'un de ses partenaires doit effectuer une action, il vérifie si il estime que ce partenaire sait qu'il doit réaliser l'action. Si ce n'est pas le cas, le robot cherche la raison de cette différence de croyance et communique à ce propos.
— lorsque le robot estime qu'un de ses partenaires pense qu'il doit effectuer une action, il vérifie si il estime également que l'action doit être réalisée. Si ce n'est pas le cas, le robot cherche la raison de cette différence de croyance et communique à ce propos.
— quand le robot s'apprête à réaliser une action, il vérifie qu'il estime que ses partenaires sont au courant de cette action, et si ce n'est pas le cas, le robot signale son action avant d'agir.
— finalement, comme l'estimation des connaissances de l'Homme par le robot peut être erronée, si le robot estime que l'Homme a toutes les connaissances pour réaliser une action mais que l'Homme n'agit pas, le robot va simplement demander à l'Homme de réaliser l'action et considérer que son estimation était erronée.

D.3.1.4 Exemple illustratif

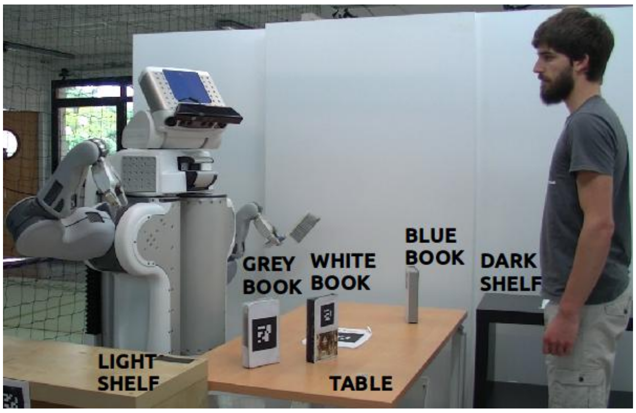


FIGURE D.7 – État du monde au début de la tâche de nettoyage de table. Le robot peut atteindre le *grey book* et le *white book* tandis que l'homme peut atteindre le *white book* et le *blue book*.

Pour illustrer les bénéfices de ce travail, nous avons utilisé une tâche ou un Homme et un robot doivent nettoyer une table ensemble. Pour cela, ils doivent enlever tous les objets initialement placés sur une table (Fig. D.7), puis le robot doit balayer la table et enfin les objets doivent être remis sur la table. Le plan initialement produit par le robot pour atteindre ce but est celui Fig. D.8.

Le robot commence à enlever le *grey book* pour le placer sur le meuble à côté de lui. Pendant ce temps, l'Homme enlève le *blue book* et s'en va (Fig. D.10(a)).

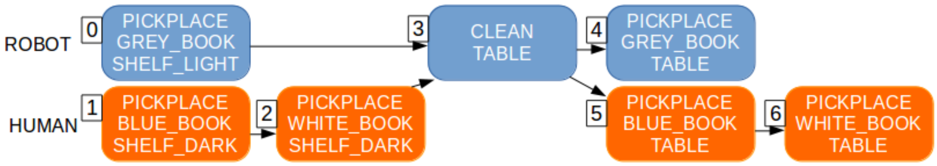


FIGURE D.8 – Plan initial pour la tâche de nettoyage de table.

Le robot termine son action. A ce stade du plan, la prochaine action qui doit être effectuée est celle de l'Homme (enlever le *white book*). Comme l'homme ne revient pas, le robot calcule un nouveau plan où il enlève le *white book* (Fig. D.9).

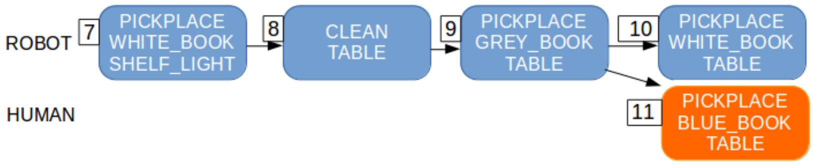


FIGURE D.9 – Deuxième plan pour la tâche de nettoyage de table.

Le robot enlève le dernier livre et balaye la table (Fig. D.10(b)). L'Homme revient alors (Fig. D.10(c)). Comme il peut voir que le *grey book* est sur le meuble à côté du robot, le robot estime que l'Homme est capable de déduire par lui même que le robot a fini sa première action (enlever le *grey book*). De même, comme l'homme peut voir le *white book*, le robot estime également que l'homme est au courant que le robot a enlevé ce livre. Cependant, l'Homme ne peut pas observer que la table a été balayée (on considère ici que la table n'était pas très sale et que l'effet de balayer la table n'est pas observable). Comme l'Homme a besoin de savoir que la table a été nettoyée pour remettre le livre qu'il avait enlevé, le robot va l'informer à propos de cette action ("J'ai balayé la table."). L'Homme a donc toutes les informations nécessaires pour finir la tâche (et aucune information superflue qu'il pouvait observer de lui même), lui et le robot finissent donc la tâche avec succès (Fig. D.10(d)).

(a) L'Homme part après avoir enlevé le premier livre.

(b) Le robot enlève le dernier livre et balaye la table.

(c) L'homme revient.

(d) L'Homme et le robot finissent la tâche avec succès.

FIGURE D.10 – Exemple illustratif d'une tâche de nettoyage de table.

D.3.2 Quand prendre les décisions pendant l'élaboration et l'exécution de plans partagés ?

D.3.2.1 Motivations

Quand plusieurs individus collaborent lors d'une action conjointe, et plus particulièrement lors de l'exécution d'un plan partagé, de nombreuses décisions doivent être prises. Certaines d'entre elles vont être implicites alors que d'autres vont nécessiter une négociation ou une adaptation entre les acteurs de l'action conjointe. Afin d'être un bon partenaire, le robot doit donc être capable de prendre les bonnes décisions au bon moment et de correctement communiquer à leur propos (ne pas devenir trop verbeux en communiquant à propos des décisions implicites tout en donnant les informations nécessaires au bon déroulement de la tâche). Nous avons identifié trois types de décisions que le robot doit prendre durant l'élaboration et l'exécution d'un plan partagé :

— **Quelles actions exécuter dans quel ordre ?** cela a été le sujet de plusieurs travaux dans le domaine de l'interaction Homme-Robot. Nous utiliserons pour gérer ce type de décisions, HATP, un planificateur de tâche capable de prendre en compte l'Homme [Lallement 2014].

— **Qui doit effectuer chaque action ?** cette décision est quelques fois implicite quand une seule personne est capable d'exécuter une action, mais peut, dans certains cas, demander une négociation ou une adaptation de la part du robot. Dans les versions précédentes d'HATP, toutes ces décisions étaient prises à l'élaboration du plan. Un des objectifs de ce travail et de reporter cette décision à l'exécution quand elle n'est pas implicite afin de gagner en fluidité et en adaptabilité par rapport à l'Homme.

— **Avec quels objets effectuer une action ?** il peut arriver que deux objets soit fonctionnellement équivalents dans le cadre de la tâche. Dans ce cas, le robot prenait auparavant à l'élaboration du plan une décision arbitraire quant à l'objet à utiliser durant une action. Le deuxième objectif de ce travail et de reporter cette décision à l'exécution quand il y a plusieurs objets fonctionnellement équivalents afin de gagner en fluidité et adaptabilité par rapport à l'Homme.

D.3.2.2 Élaboration du plan partagé

Afin de pouvoir reporter certaines décisions à l'exécution, nous avons effectué deux changements à la façon dont HATP élabore un plan :

— Afin de reporter la décision de qui doit effectuer une action quand plusieurs agents peuvent effectuer cette action, nous avons introduit dans HATP un agent virtuel appelé *l'agent X*. Cette agent aura comme capacités l'intersection des capacités de l'Homme et du robot et aura un coût bien plus faible que les autres agents quand il réalisera une action. De cette manière, quand une action pourra être réalisée soit par l'Homme, soit par le robot, elle sera automatiquement attribuée à *l'agent X* et la décision sera prise à l'exécution.

— Nous avons également introduit ce que l'on a appelé des *objets haut niveaux*. Ces *objets haut niveaux* seront utilisés dans le plan par HATP quand deux objets fonctionnellement équivalents pourront être utilisés pour réaliser une action (par exemple CUBE_ ROUGE à la place de CUBE_ROUGE1 ou CUBE_ROUGE2).

D.3.2.3 Exécution du plan partagé

Une fois le plan élaboré, le robot doit être capable de l'exécuter en prenant les bonnes décisions au bon moment. Pour cela, grâce aux travaux antérieurs à cette thèse [Fiore 2014], le robot est capable de maintenir le plan partagé et de réagir aux actions inattendues de l'Homme. Quand le robot aura à effectuer une action du plan, il choisira en priorité une action qui lui est allouée par rapport à une action allouée à *l'agent X* de manière à laisser le choix le plus longtemps possible à l'Homme d'effectuer cette action ou non.

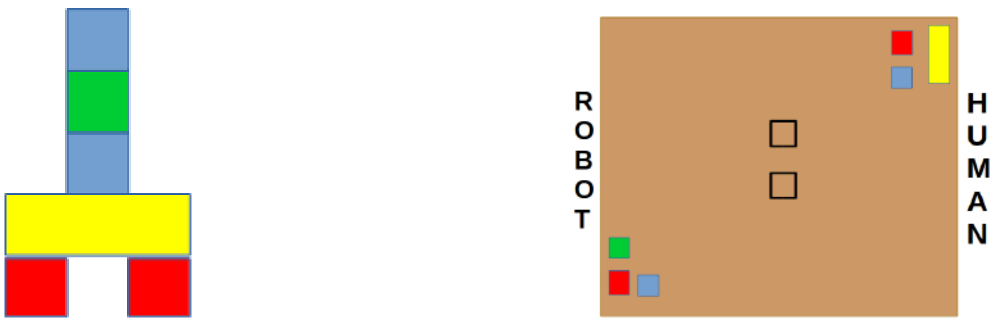
Quand le robot devra choisir qui doit réaliser une action allouée à *l'agent X*, il vérifiera dans un premier temps quels sont les agents réellement disponibles pour effectuer cette action. Si le robot est le seul à pouvoir réaliser l'action (e.g. l'homme est déjà en train d'effectuer une autre action), il prendra l'initiative de réaliser l'action. Si l'Homme et le robot peuvent tous les deux réaliser l'action, le robot aura alors deux différents modes possibles :

— **Le mode négociation :** où le robot demande à l'homme si il veut réaliser l'action et agit (ou non) en fonction de sa réponse.
— **Le mode adaptation :** où le robot attend un temps court de voir si l'homme prend l'initiative de réaliser l'action, et, si ce n'est pas le cas, prend lui même l'initiative de la réaliser.

Une fois une action allouée par le robot, il calcule un nouveau plan pour vérifier que cette allocation n'a pas d'autres implications dans le plan.

Finalement, quand le robot doit réaliser une action comportant un *objet haut niveau*, le robot va reporter la décision au dernier moment possible pour laisser plus de latitude à l'Homme. Pour prendre la décision de quel objet utiliser, le robot pourra utiliser des coûts prenant en compte l'Homme tels que la distance entre les agents et les différents objets. L'exécution de l'action par le robot se fera en boucle fermée et en surveillant l'activité de l'Homme de manière à pouvoir changer de décision si l'Homme prend une initiative en conflit avec la décision précédente.

D.3.2.4 Exemple illustratif



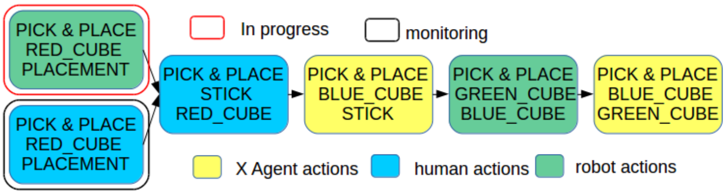
(a) But de la tâche (vue de côté).

(b) Un possible état de départ (vue de haut).)

FIGURE D.11 – Description de la tâche de construction de blocs. L'Homme et le robot doivent construire une pile ensemble.

Pour illustrer les bénéfices de ce travail, nous avons utilisé une tâche inspirée de celle présentée dans [Clodic 2014]. Un Homme et un robot doivent réaliser une construction avec des blocs colorés comme représenté Fig. D.11(a). Au début de la tâche, l'Homme et le robot ont chacun plusieurs blocs de couleur à leur disposition comme par exemple Fig. D.11(b). Deux emplacements identiques sont placés au centre de la table pour indiquer ou mettre les deux premiers cubes rouges.

Le plan produit initialement par HATP pour réaliser la tâche peut être trouvé Fig. D.12(a). Le robot attrape le cube rouge à sa disposition et choisit de le placer sur l'emplacement à sa droite (Fig. D.12(b)). Cependant, l'Homme prend son cube rouge et choisit de le placer sur le même emplacement que celui choisi par le robot (Fig. D.12(c)). Le robot interrompt son action et s'adapte en plaçant son cube sur l'autre emplacement (Fig. D.12(d)). L'homme pose alors le bâton jaune sur les cubes rouges. Dans cet exemple, le robot utilise le mode **négociation** pour choisir

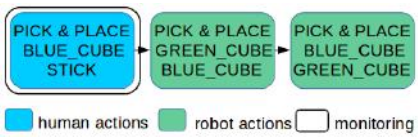


(a) Plan initial.

(b) Le robot choisit de mettre son cube rouge sur l'emplacement à sa droite.

(c) L'Homme pose son cube sur l'emplacement que le robot a choisit.

(d) Le robot s'adapte en changeant son choix d'emplacement.



(e) Deuxième plan.

FIGURE D.12 – Exemple illustratif d'une tâche de construction de blocs.

qui va mettre le premier cube bleu. Le robot demande donc à l'Homme si il veut poser le cube bleu ("Voulez-vous poser le cube bleu ?"). L'Homme répond oui, le robot calcule donc un nouveau plan ou il posera le second cube bleu Fig. D.12(e). Finalement, l'Homme et le robot effectuent leurs dernières actions et réalisent la tâche avec succès.

D.3.3 Évaluation du système

D.3.3.1 Tâche et conditions

Afin d'évaluer le nouveau système développé concernant la gestion des plans partagés par le robot, nous avons développé une tâche d'inventaire. Dans cette tâche, l'Homme et le robot doivent scanner différents cubes de couleur et les ranger dans une boîte de même couleur. Au début de la tâche chaque agent a une pile de cubes de différentes couleurs à laquelle seulement lui peut accéder. Ces piles contiennent des cubes bleus, verts et rouges. Pour que les cubes soit scannés, les agents doivent les poser un par un sur une des deux zones de scan sur la table devant le robot (voir Fig. D.13). Une fois un cube sur une zone, le robot peut le scanner en orientant sa tête vers le cube et en allumant une lumière rouge en direction du cube (voir Fig. D.15). Une fois un cube scanné, il peut être rangé dans une boite de la même couleur. Le robot a accès à une boite bleue et à une boite rouge tandis que l'Homme a accès à une boite verte et à une boite rouge (voir Fig. D.13). Cette tâche comporte deux particularités intéressantes pour notre système :

— Comme la pile de l'homme et ses boîtes sont situées dans des pièces différentes que celle du robot (voir Fig. D.13), si le robot scanne un cube quand l'Homme est parti chercher ou ranger un cube, l'Homme ne pourra pas savoir que le cube a été scanné sauf si le robot l'en informe (pas d'effets visibles).

— La répartition des boites entre les agents fait que seul le robot peut ranger les cubes bleus, seul l'Homme peut ranger les cubes verts mais qu'ils peuvent tous les deux ranger les cubes rouges.

Nous avons évalué notre système en simulation et lors d'une étude utilisateur. Pour faire cela, nous avons comparé 4 différentes conditions :

— avec le système original, appelé **système de référence (RS)**, où toutes les décisions du robot sont prises durant l'élaboration du plan et sans estimation des états mentaux de l'Homme :
  — **RS-NONE mode :** le robot ne verbalise rien (sauf en cas de stricte nécessité).
  — **RS-ALL mode :** le robot informe à propos de toutes les actions qu'il doit faire et que l'homme doit réaliser ainsi qu'à propos de toutes les actions manquées par l'Homme.

— avec le nouveau système développé présenté précédemment (**NS**) :
  — **NS-N mode :** le robot utilise le mode **négociation** pour prendre une décision concernant les actions de *l'agent X*.
  — **NS-A mode :** le robot utilise le mode **adaptation** pour prendre une décision concernant les actions de *l'agent X*.

D.3.3.2 Évaluation en simulation

Pour évaluer notre système en simulation, nous avons fait tourner la tâche avec différents états de départ ou les piles des agents étaient aléatoirement composées. Durant ces simulations, le robot était confronté à plusieurs types d'Hommes simu-

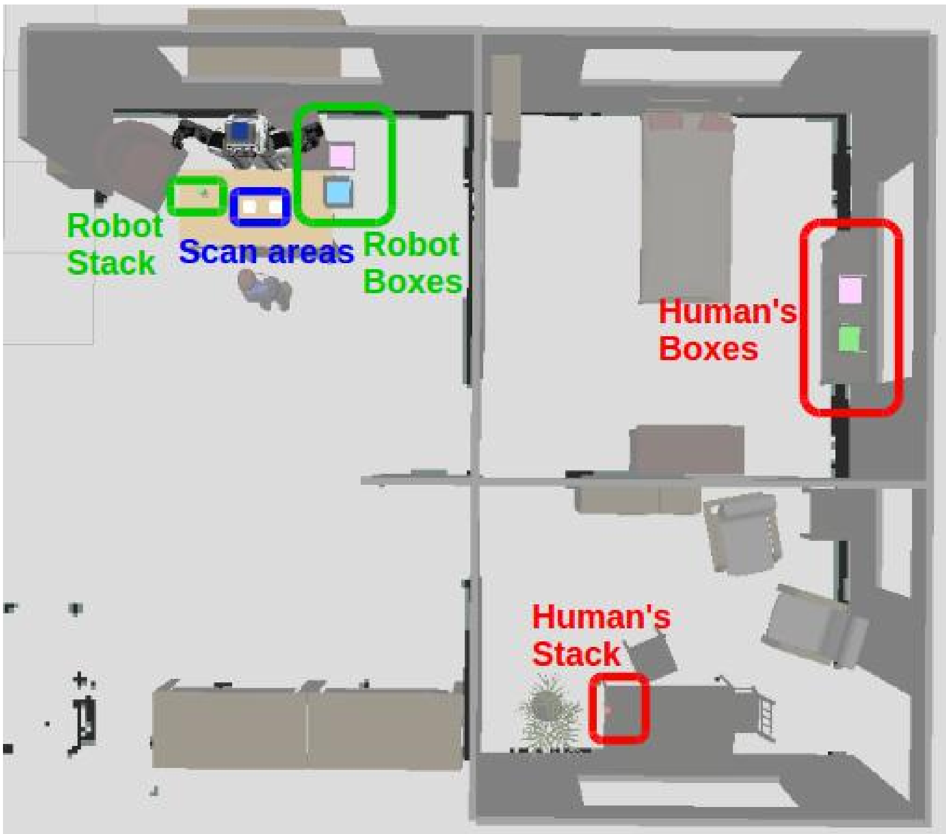


FIGURE D.13 – Etat initial pour la tâche d'inventaire. L'Homme et le robot doivent prendre les cubes de leur pile pour les mettre sur une zone de scan. Puis, le robot doit scanner le cube et enfin chaque cube doit être rangé dans une boite de même couleur. L'Homme a accès à une boite verte et à une boite rouge tandis que le robot a accès à une boite bleue et à une boite rouge.

lés :

— *l'Homme "aimable" (cas **K**) :* qui adapte son comportement à ce que verbalise le robot. Concernant les cubes rouges, il peut choisir de ne jamais les ranger (**lazy-K**), les ranger systématiquement (**hurry-K**) ou les ranger avec une probabilité de 50% (**50%-K**)

— *l'Homme "têtu" (cas **S**) :* qui n'adapte pas son comportement à ce que verbalise le robot. Concernant les cubes rouges, il peut choisir de ne jamais les ranger (**lazy-S**), les ranger systématiquement (**hurry-S**) ou les ranger avec une probabilité de 50% (**50%-S**)

Dans tous les cas, l'Homme participe activement au dépôt des cubes de sa pile sur les zones de scan et au rangement des cubes verts et répond aux questions posées par le robot.

Les données mesurées durant ces simulations sont :

— *le nombre d'interactions verbales :* entre l'Homme et le robot (information donnée par le robot ou question posée), Tab. D.1.

— *le nombre de décisions incompatibles :* les deux acteurs prennent la même décision concernant une action (les deux essayent de la réaliser ou de ne pas la réaliser), Tab. D.2.
— *le temps d'exécution total :* pour réaliser la tâche, Fig. D.14.

	RS-NONE	**RS-ALL**	**NS-N**	**NS-A**
50%-K	2.4 (0.84)	20.7 (1.34)	3.4 (1.51)	2 (1.33)
hurry-K	1.8 (0.79)	21.1 (2.08)	1.9 (1.10)	2.2 (1.13)
lazy-K	3.0 (1.33)	21 (1.56)	3.3 (1.42)	1.6 (1.17)
50%-S	2.5 (1.43)	23.9 (1.59)	3.3 (1.49)	1.7 (0.95)
hurry-S	1.5 (0.97)	20.9 (1.29)	2.4 (1.89)	1.9 (0.99)
lazy-S	3.2 (0.92)	25.2 (1.55)	2.8 (1.68)	1.8 (1.14)

TABLE D.1 – **Nombre d'interactions verbales** : questions posées par le robot dans le mode négociation et nombre d'informations verbalisées. Ces résultats correspondent à la moyenne sur 10 essais et leur déviation standard associée.

Plusieurs choses peuvent être observées par rapport aux résultats obtenus :
— le mode **RS-NONE** est celui comportant le plus de décisions incompatibles (du fait que le robot ne communique et de s'adapte pas par rapport aux cubes rouges et aux choix des zones de scan). Ce mode comporte également les plus grands temps d'exécution, plus spécialement dans le cas de l'Homme "têtu" car le robot perd du temps à attendre qu'il exécute des actions qu'il ne veut pas réaliser ou à attendre que l'Homme range un cube dont il ne sait pas qu'il a été scanné.
— comme attendu, le mode **RS-ALL** est celui avec le plus d'interactions verbales. Cependant, ces interactions verbales ne suffisent pas à supprimer toutes les décisions incompatibles, surtout dans le cas de l'homme "têtu" ou le temps d'exécution est également plus élevé.
— on peut voir que les performances du **nouveau système** sont globalement meilleures que celles de l'ancien système. En effet pour beaucoup moins d'informations verbalisées, il permet de supprimer les décisions incompatibles et de réduire le temps d'exécution dans le cas de l'Homme "têtu". Le mode **adaptation** obtient les mêmes résultats que le mode **négociation** mais avec moins d'interactions verbales.

D.3.3.3 Étude utilisateur

Adaptation de la tâche pour l'étude utilisateur : Avant de réaliser l'étude utilisateur, nous avons effectué quelques pré-tests qui nous ont permis d'identifier plusieurs problèmes et d'y remédier avec des adaptations de la tâche :
— **introduction d'une cassette rouge :** afin d'assurer qu'il y ait forcement une prise de décision par rapport à un objet rouge (de temps en temps la configuration faisait qu'aucune décision n'était nécessaire), nous avons ajouté une cassette rouge qui doit être scannée et rangée une fois que tous les cubes

	RS-NONE	RS-ALL	NS-N	NS-A
50%-K	2.9 (0.99)	0.9 (0.57)	0.6 (0.7)	0.3 (0.48)
hurry-K	2.5 (0.97)	1.0 (0.94)	0.6 (0.52)	0.4 (0.52)
lazy-K	3.5 (1.08)	0.8 (0.63)	0.5 (0.7)	0.5 (0.53)
50%-S	2.9 (1.45)	1.9 (0.99)	0.6 (0.52)	0.5 (0.97)
hurry-S	2.3 (1.34)	1.0 (0.82)	0.5 (0.53)	0.4 (0.52)
lazy-S	3.5 (0.97)	2.6 (1.84)	0.3 (0.67)	0.4 (0.52)

TABLE D.2 – **Nombre de décisions incompatibles** : les deux acteurs prennent la même décision concernant une action (les deux essayent de la réaliser ou de ne pas la réaliser). Ces résultats correspondent à la moyenne sur 10 essais et leur déviation standard associée.

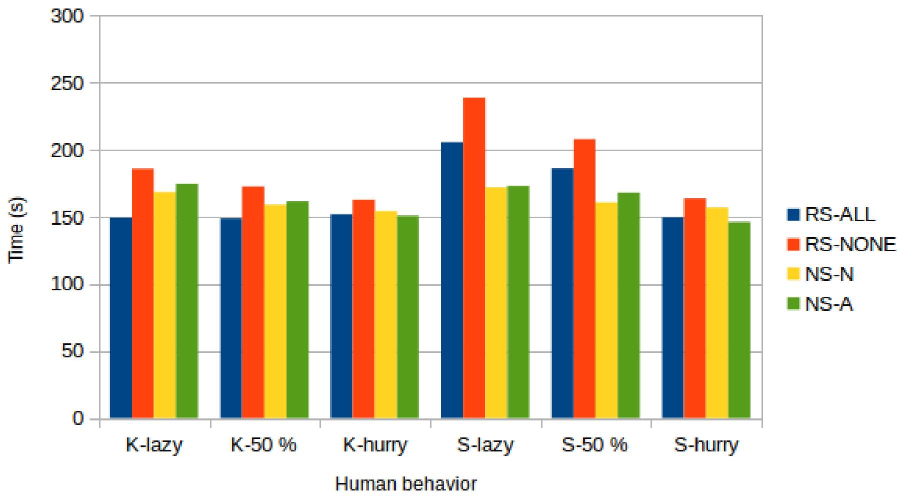


FIGURE D.14 – Temps en secondes nécessité pour chaque système pour réaliser la tâche dans chaque condition (moyennes sur 10 essais).

ont été rangés. L'Homme et le robot ont chacun initialement une cassette rouge mais une seule doit être scannée et rangée.

— **Tâche de distraction :** afin d'être surs d'obtenir un manque de connaissance à un moment de la tâche (de temps en temps l'homme ne s'absentait jamais lors d'une action de scan), nous avons rajouté une tâche de construction avec des Légos pour le sujet à un moment de la tâche dans un lieu ou il ne peut pas voir le robot.

Questionnaire et protocole : 21 sujets (8 femmes et 13 hommes) ont interagi avec le robot pour réaliser la tâche dans les quatre conditions décrites précédemment. L'ordre de ces conditions et les compositions des piles des agents étaient aléatoires. A leur arrivée, les participants étaient introduits à l'environnement de travail et au robot par l'expérimentateur. Ensuite, les participants avaient à lire

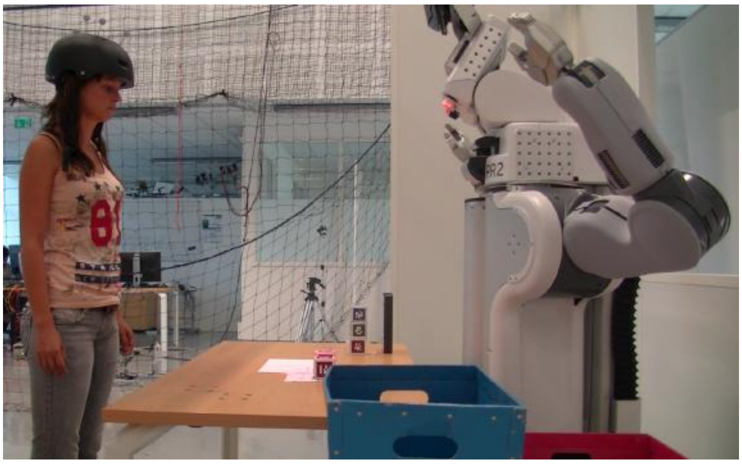


Figure D.15 – Le robot PR2 interagissant avec un sujet pour réaliser la tâche. Le robot scanne un objet avant de le ranger.

les consignes de la tâche et l'expérimentateur vérifiait leur bonne compréhension. Les participants réalisaient une rapide tâche de familiarisation avant de réaliser la vraie tâche. Après chaque interaction avec le robot (pour chaque condition), les participants avaient à remplir un questionnaire leur permettant d'évaluer le comportement du robot. Comme nous n'avons pas trouvé dans la littérature existante de questionnaires permettant d'évaluer la prise de décision haut niveau d'un robot lors d'une tâche de collaboration avec l'Homme, nous avons conçu ce questionnaire en nous basant sur le modèle d'expérience utilisateur UX [Mahlke 2008] et en ajoutant des dimensions spécifiques à la prise de décision. Ce questionnaire est composé de plusieurs dimensions :

— **Dimension de collaboration**, basée sur [Weistroffer 2014] et permettant d'évaluer la perception de l'utilité et de l'utilisabilité du robot [Davis 1989].
— **Dimension d'interaction**, basée sur [Lallemand 2015] et permettant d'évaluer l'intention d'utilisation [Davis 1989].
— **Dimension de perception du robot**, basée sur le questionnaire Godspeed [Bartneck 2009] et permettant d'évaluer comment le sujet perçoit le robot en général [Hassenzahl 2003].
— **Dimension émotions,** reprise de l'AffectButton [Broekens 2013] et permettant d'évaluer les émotions du sujet lors de l'interaction.
— **Dimension verbale** permettant d'évaluer comment le sujet a perçu les interactions verbales avec le robot.
— **Dimension d'action** permettant d'évaluer comment le sujet a perçu la prise de décision du robot par rapport au choix d'exécution des actions.

Ces dimensions étaient évaluées grâce à des questions où le sujet devait se placer sur une échelle de 100, sauf pour la dimension émotion ou le sujet avait à choisir entre plusieurs smileys.

Hypothèses : Nous avons émis plusieurs hypothèses avant l'étude :
— **Hypothèse 1 :** le nouveau système sera préféré par les utilisateurs à l'ancien système.
— **Hypothèse 2 :** concernant le nouveau système, au contraire des résultats de simulation, le mode négociation sera préféré par les utilisateurs au mode adaptation.

Résultats : La cohérence interne de questionnaire a été vérifiée à la suite de cette étude (alpha de Cronbach supérieur à 0,7 pour toutes les dimensions du questionnaire). Concernant les scores des différentes conditions, les résultats totaux du questionnaire peuvent être trouvés en Fig D.16.

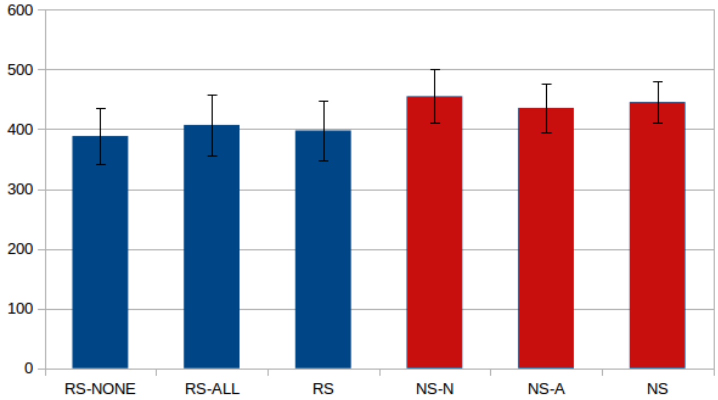


FIGURE D.16 – Scores totaux du questionnaire utilisateur. Addition des scores de toutes les dimensions précédemment remis sur une échelle de 100.

Les scores du nouveau système (NS) ont été trouvés significativement plus élevés ($p < 0.05$) que ceux de l'ancien système (RS). Cela permet donc de vérifier la première hypothèse comme quoi le nouveau système a été préféré des sujets. Concernant le nouveau système, si l'on regarde la dimension verbale du questionnaire, le score de la condition **négociation** a été trouvé significativement plus élevé que celui de la condition **adaptation** ($p < 0.05$). Il n'y a pas eu de différence significative concernant les autres dimensions. Comme la seule différence entre ces deux modes consiste à poser ou non une question quand il y avait une décision à prendre concernant un objet rouge (comportement verbal), nous pouvons donc considérer la seconde hypothèse comme validée également.

D.3.3.4 Conclusion

L'évaluation du système en simulation et lors d'une étude utilisateur a permis de montrer que le nouveau système développé a de meilleures performances et une meilleure appréciation par l'utilisateur que l'ancien système. Concernant les deux modes possible du nouveau système, des utilisateurs naïfs comme ceux de l'étude utilisateur préfère le mode **négociation**. Cependant, pour des utilisateurs plus experts, le mode **adaptation** a montré de meilleurs résultats en simulation. L'étude utilisateur nous a également permis de développer et valider un questionnaire permettant d'évaluer la prise de décision haut niveau du robot lors d'une tâche de collaboration avec l'Homme.

D.4 Autres contributions à l'action conjointe Homme-Robot

D.4.1 Communication non-verbale : qu'est ce que le robot doit faire avec sa tête ?

D.4.1.1 Motivations et précédents travaux

Pour communiquer entre eux quant ils collaborent sans être trop verbeux, les hommes utilisent fréquemment la communication non-verbale [Ekman 1969, DePaulo 1992]. Durant l'action conjointe Homme-Robot, le robot doit également être capable de communiquer à son partenaire toutes les informations dont il a besoin sans être trop intrusif. Pour cela, il doit donc être capable d'avoir un comportement non-verbal adapté à l'action conjointe. La communication non-verbale vient de multiples sources (expression faciales [LaBarre 1947], postures [Mehrabian 1969], regard [Mutlu 2009a], etc...). Dans cette thèse, nous nous sommes concentré sur l'utilisation de la tête du robot, remplaçant les signaux donnés par le regard en l'absence de pupilles pour le robot [Imai 2002].

Les sciences sociales ont permis de déterminer plusieurs utilisations du regard durant l'action conjointe entre Hommes :

— **Aide au dialogue et à la prise de tour :** le regard est très utilisé lors du dialogue [Argyle 1976] et plus spécifiquement pour signaler les changements d'orateur [Kendon 1967].

— **Aide à la compréhension des actions :** les acteurs d'une action conjointe vont agir différemment que lorsqu'ils agissent seul [Becchio 2010, Vesper 2010]. Plus spécifiquement, l'utilisation du regard lors d'une action va permettre aux partenaires de mieux interpréter les intentions de l'acteur [Castiello 2003, Pierno 2006].

— **Aide à la compréhension des états mentaux :** l'observation du regard du partenaire permet de mieux prendre sa perspective afin de mieux estimer ses connaissances [Furlanetto 2013].

En robotique, plusieurs études ont montré l'intérêt du comportement non-verbal du robot [Furlanetto 2013, Häring 2012]. Beaucoup de travaux se sont concentrés sur l'utilisation de la tête lors du dialogue [Mutlu 2009b, Boucher 2010, Skantze 2014]. Seuls quelques uns portent sur l'utilisation du regard durant l'action conjointe et ont montré qu'un bon comportement de la part du robot aide à la coordination lors de l'exécution du plan partagé [Lallée 2013] et à la prise de décision de l'Homme [Boucher 2012].

D.4.1.2 Réflexion concernant les signaux et comportements nécessaires

Sur la base d'une étude bibliographique des comportements humains et des travaux sur la tête du robot, nous avons identifié ce que nous pensons être des composants nécessaires à un bon comportement de la tête du robot lors de l'action conjointe :

(a) Début de la tâche. Le cube bleu et le cube vert sont accessibles par l'Homme et les cubes noir et rouge sont accessibles par le robot.

(b) Fin de la tâche. La pile doit être construite dans un ordre précis (rouge, noir, bleu, vert).

FIGURE D.17 – Tâche utilisée dans l'étude utilisateur en ligne. Dans cette tâche, l'Homme et le robot doivent construire une pile de cubes colorés.

— **Lorsque le robot agit :** lors de ses actions, le robot doit utiliser sa tête à la fois pour la bonne réalisation de l'action d'un point de vue fonctionnel (présence de caméras à l'intérieur de la tête) mais également pour indiquer à ses partenaires ce qu'il fait et ce qu'il va faire ensuite.
— **Lorsque le robot parle :** lors d'un dialogue, il est important pour le robot de regarder l'homme au bon moment ainsi que les objets dont il parle.
— **Le robot observe :** le robot doit se servir de sa tête pour montrer son intérêt et sa compréhension des actions de l'Homme.
— **Le robot se coordonne :** le robot doit fournir les signaux appropriés et nécessaires au bon déroulement du plan partagé.

D.4.1.3 Étude approfondie de certains signaux

Nous avons étudié dans plus de détails certains composants des comportements définis précédemment. Pour faire cela, nous avons mené une étude utilisateur en ligne à base de vidéos. Dans cette étude, nous avons demandé à 59 personnes (30 femmes et 29 hommes) de regarder plusieurs vidéos courtes ou le comportement de la tête du robot changeait et d'évaluer ces comportements grâce à un petit questionnaire. Dans ces vidéos, l'Homme et le robot avaient à construire une pile de cubes colorés (comme illustré Fig. D.17).

Les comportements testés et les résultats obtenus sont les suivants :

— **Anticipation des actions du robot :** nous avons comparé un comportement du robot ou il anticipait sa prochaine action avec sa tête (il regarde le cube à prendre avant de commencer son action) au même comportement sans cette anticipation. Nous n'avons pas trouvé de différence significative entre ces deux comportements. En effet, certains des sujets étaient perturbés par le fait que le robot regarde le cube avant d'agir et d'autres n'ont pas vu la différence. Une possible explication pour cela est que la tâche ne demandait pas d'anticipation de la part du robot car les deux participants savaient quelle action était nécessaire à chaque moment (ordre de la pile prédéfini).

— **Suivre l'activité de l'homme :** nous avons comparé différents moyens pour le robot de suivre avec sa tête l'activité de l'Homme. Dans la première condition, le robot regardait la main de l'homme dès qu'elle était en mouvement et la tête sinon. Dans la seconde, le robot regardait la main de l'Homme quand elle était dans une zone de travail définie au dessus de la table et la tête sinon. Finalement, dans la dernière condition, le robot regardait la main de l'Homme quand elle était en mouvement et dans la zone de travail et la tête sinon. Cette dernière condition a été significativement préférée aux deux autres par les sujets.

— **Montrer la compréhension des actions de l'Homme :** nous avons comparé un comportement ou le robot marquait un arrêt avec sa tête quand l'Homme réalisait une action de manière à montrer qu'il avait détecté l'action à une condition sans cet arrêt. Nous n'avons pas trouvé de différence significative entre ces deux conditions, les sujets ayant du mal à trouver les différences entre les deux vidéos.

— **Gérer l'inaction de l'Homme :** dans ces vidéos, l'Homme mettait du temps à réaliser une de ses actions. Nous avons comparé trois différentes réaction de la part du robot. Dans la première, le robot ne changeait pas son comportement de base face à cette inaction. Dans la seconde, le robot donnait un signal à l'Homme avec sa tête en regardant le cube que l'Homme devait prendre. Dans la dernière, le robot donnait un signal similaire mais cette fois ci regardait le cube de l'Homme puis la pile. Les deux conditions ou le robot donnait un signal à l'Homme ont été notées significativement mieux que celle sans signal montrant l'importance du signal du robot. Aucune différence n'a été trouvée entre les deux différents signaux.

— **Aide à la prise de tour :** nous avons comparé différentes manières pour le robot de gérer le changement d'acteur dans la tâche (passage d'une action du robot à une action de l'Homme). Dans deux conditions, le robot ne donnait pas de signal particulier à l'Homme. Il regardait simplement l'Homme soit à la fin de son action dans une condition, soit après s'être retiré de son action dans l'autre condition. Dans les deux autres conditions, le robot regardait le cube que l'Homme devait poser avant de regarder l'Homme. Comme pour les deux précédentes conditions, le robot faisait cela à la fin de son action dans une condition et après s'être retiré dans l'autre condition. Les deux conditions où le robot regardait le cube de l'Homme ont été trouvées significativement meilleures que les deux autres, montrant l'intérêt du signal du robot. Aucune différence n'a été trouvée concernant le timing du signal (avant ou après le retrait).

— **Choisir un objet d'attention :** dans le dernier scénario, l'Homme commençait à prendre un cube pendant que le robot était toujours en train de poser le sien. Dans une condition, le robot continuait son action sans regarder l'Homme. Dans la seconde condition, le robot regardait l'Homme mais sans interrompre sa propre action. Dans la dernière condition, le robot interrompait son action pour regarder celle de l'Homme. La condition ou le robot ne regarde pas l'Homme a été trouvée significativement moins bonne que les deux autres. Aucune différence n'a été trouvée entre les deux autres conditions. Cela montre l'importance pour le robot de regarder l'action de l'Homme même si il doit pour cela interrompre sa propre action.

D.4.1.4 Proposition d'architecture pour le comportement de la tête du robot

A partir de l'étude bibliographique et des résultats de l'étude présentée précédemment, nous avons proposé une architecture pour gérer la tête du robot. Cette architecture peut être trouvée Fig. D.18.

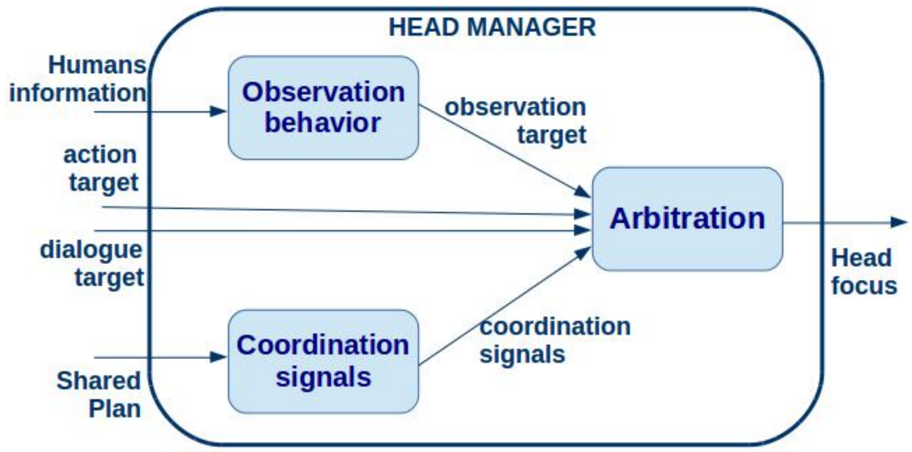


FIGURE D.18 – Architecture pour gérer la tête du robot. Un module d'arbitrage choisit où le robot doit regarder en se basant sur plusieurs comportements et signaux produits par des modules en amont.

Un premier module permet au robot de générer un comportement pour montrer sa compréhension de l'activité de l'Homme. Basé sur les résultats de l'étude précédente, nous proposons d'implémenter un comportement où le robot regarde successivement la tête et la main de l'Homme en se basant sur le mouvement de la main et sa position (dans ou en dehors de zones de travail). Nous proposons également d'implémenter un comportement qui permet au robot de répondre aux regards de l'Homme (le robot regarde l'Homme si il le regarde et regarde l'objet de l'attention de l'Homme si l'Homme regarde fixement un objet).

En entrée de l'architecture proposée, nous trouvons des points d'intérêt venant du module d'exécution d'action du robot et du module de dialogue. Ces deux modules fournissent tous les deux l'objet ou la personne la plus pertinente à regarder en fonction de l'action du robot et de la conversation en cours.

Un autre module permet de créer des signaux à donner à l'Homme concernant l'exécution du plan partagé. Nous proposons d'implémenter les deux signaux étudiés précédemment (signal quand l'Homme n'agit pas et signal d'aide à la prise de tour).

Finalement, un module d'arbitrage permet de choisir entre les différents comportements et signaux générés par les autres modules en se basant sur ce que fait le robot et les différentes priorités des signaux et comportements.

D.4.2 Combiner apprentissage et planification

D.4.2.1 Motivations et travaux précédents

Concernant la prise de décision en robotique, on retrouve deux grandes écoles de pensée qui ont chacune leurs avantages et désavantages : l'apprentissage et les processus déterministes (ou planification). L'apprentissage est généralement "peu coûteux" au sens où une décision est prise rapidement et une solution va toujours être proposée quelque soit le problème. Cependant, la phase d'apprentissage requière une grande quantité de données et/ou une longue période d'apprentissage durant laquelle le robot va produire des comportements inconsistants et perturbants pour l'utilisateur humain. La planification va être plus lente à prendre une décision, particulièrement dans le cas d'un environnement ou d'une tâche complexe, mais va pouvoir prendre en compte des règles sociales et assurer la validité de la solution proposée dans son ensemble. L'idée de ce travail est de combiner ces deux techniques dans le contexte de le prise de décision pour l'interaction Homme-Robot.

Ces deux écoles de pensées sont inspirées des différents comportements des mammifères et de l'Homme : le comportement *dirigé vers un but* pour la planification et le comportement *habituel* pour l'apprentissage [Dickinson 1985]. Différentes études ont été menées en neuroscience pour trouver comment alterner entre ces comportements [Pezzulo 2013, Lesaint 2014, Viejo 2015]. En robotique, de nombreux travaux ont été réalisés en planification [Ingrand 2014] et en apprentissage [Kober 2011, Martins 2010, Stulp 2013] mais peu d'entre eux se concentrent sur comment combiner ces approches.

D.4.2.2 Présentations des différents experts

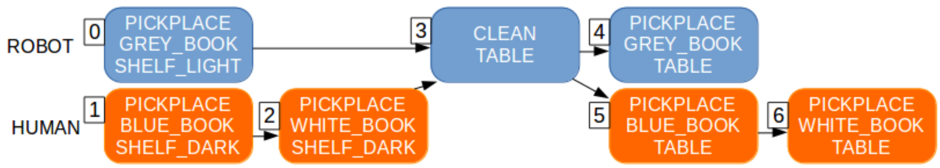


FIGURE D.19 – Un exemple de plan produit par HATP. Ce plan permet à un Homme et à un robot de nettoyer une table en enlevant tous les objets dessus, la nettoyant puis replaçant tous les objets enlevés précédemment.

Dans le travail présenté dans cette thèse, deux experts ont été utilisés pour modéliser les deux comportements évoqués précédemment :

— Le comportement *dirigé vers un but* est fourni par HATP [Lallement 2014], un planificateur HTN conçu pour le contexte de l'interaction Homme-robot. HATP prend en compte les préconditions et effets des différentes actions possibles pour construire un plan qui permet d'atteindre un but précis depuis un contexte donné (e.g. Fig. D.19). HATP permet de calculer un plan complet qui permet d'atteindre un but donné et qui prend en compte des coûts concernant l'Homme. Cependant, il ne permettra pas d'apprendre du comportement de l'Homme en direct, les coûts étant codés à l'avance. Son temps de décision sera plus lent que celui de l'autre expert mais ne nécessite pas de période d'apprentissage.

— Le comportement *habituel* est produit par un algorithme d'apprentissage par renforcement sans modèle [Renaudo 2014] qui permet d'apprendre une action à exécuter pour chaque état possible en se basant sur un principe de récompense. Cet algorithme est implémenté comme un réseau neuronal (voir Fig. D.20). Cet algorithme permet de toujours proposer rapidement une action à exécuter par le robot. Cependant une longue phase d'apprentissage est nécessaire pendant laquelle le robot aura un comportement inconsistant au début et à chaque changement de la tâche.

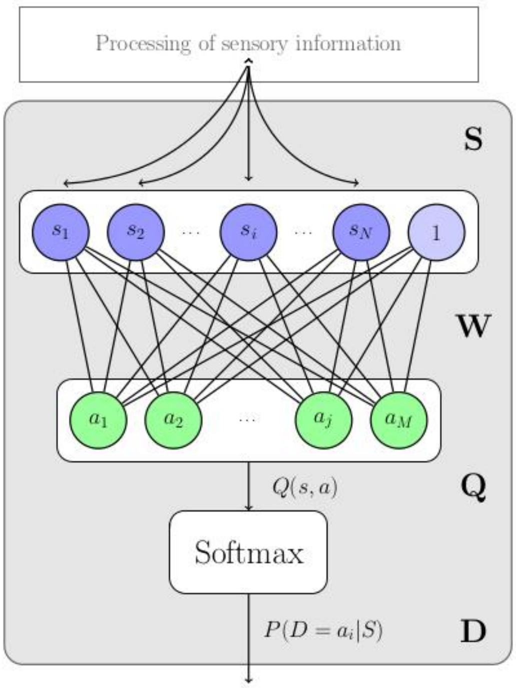


Figure D.20 – L'expert du comportement habituel est un algorithme d'apprentissage implémenté comme un réseau neuronal. Il reçoit en entrée un état S qui est projeté sur un neurone d'entrée s_i définissant une activité d'entrée. L'activité est propagée grâce aux poids du réseau neuronal W pour générer une activité sur le niveau d'action. Cette activité correspond à la valeur $Q(S, a_j)$ et est convertie en une probabilité de distribution permettant à l'expert de prendre une décision D sur la prochaine action à exécuter.

D.4.2.3 Première architecture : une preuve de concept

Architecture : La première architecture développée pour combiner les deux experts peut être trouvée Fig. D.21. Dans cette architecture les deux experts sont placés en parallèle. Le module d'évaluation de la situation prend les données de la perception et maintient l'état du monde courant. Cet état du monde est utilisé par la supervision pour calculer la récompense et par les experts pour prendre une décision. Les propositions des deux experts sont envoyées au méta-contrôleur qui décide de l'action à exécuter (de manière aléatoire). Le superviseur exécute l'action avec l'aide des modules de plus bas niveau.

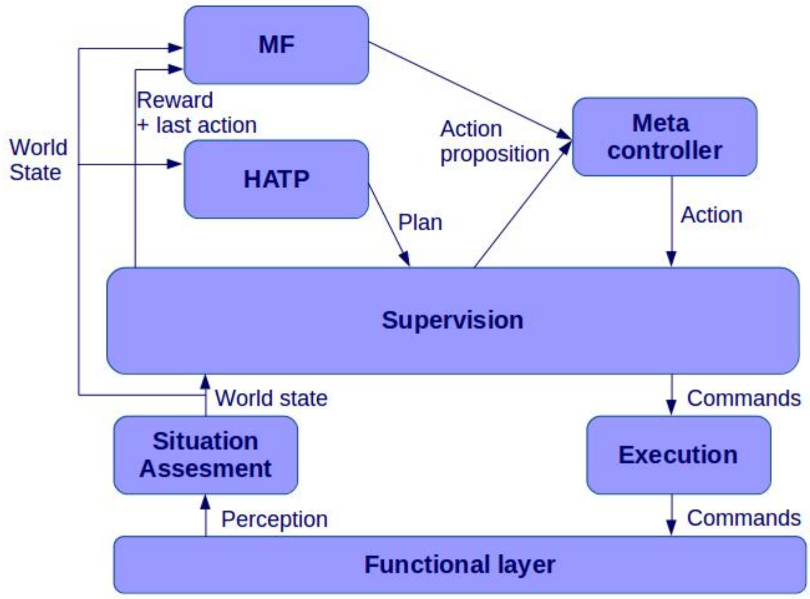
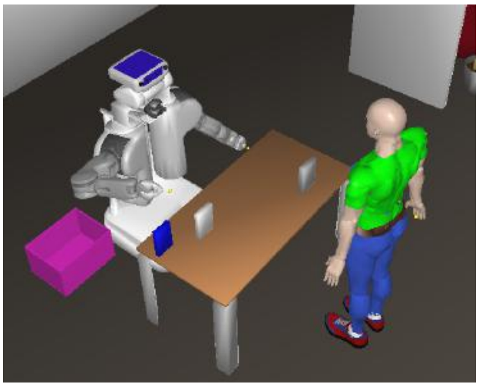


Figure D.21 – Première architecture développée pour combiner les deux experts.

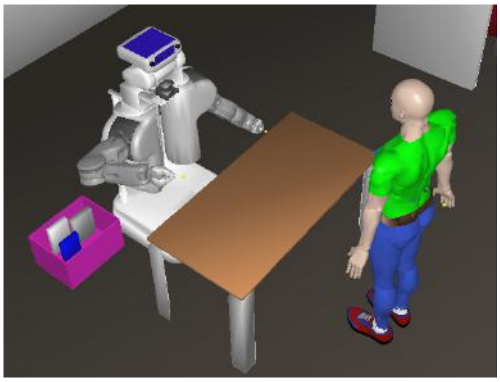
Tâche : Nous avons testé cette architecture sur une tâche simple en simulation illustrée Fig. D.22. Dans cette tâche, l'Homme et le robot doivent enlever des objets d'une table et les mettre dans une boite rose. Au début de l'interaction, deux objets sont accessibles uniquement par le robot et un autre uniquement par l'Homme. La boite est accessible uniquement par le robot. Pour réaliser la tâche, l'Homme et le robot peuvent exécuter différentes actions (prendre un objet, ranger un objet, s'échanger un objet ou attendre).

Le comportement de l'Homme est simulé dans cette expérience. L'Homme est collaboratif : il exécute toutes les actions prévues pour lui dans HATP et participe à tous les échanges d'objets entrepris par le robot.

Résultats : Pour tester notre architecture, nous avons réalisé la tâche avec chaque expert seul dans un premier temps puis avec la combinaison des deux. La tâche a été



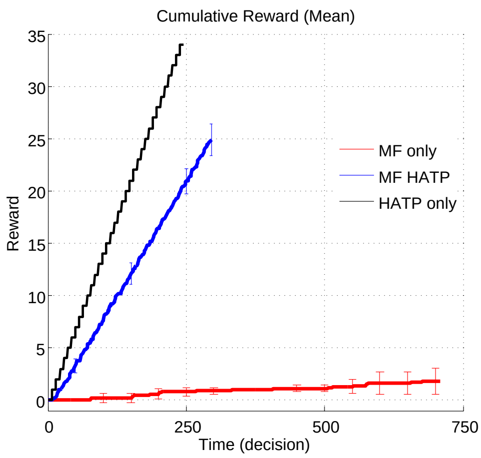
(a) Situation initiale



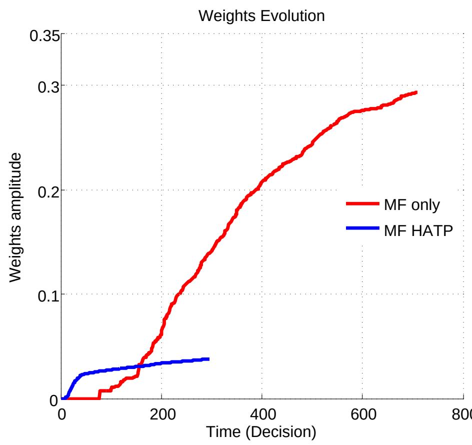
(b) Situation finale

FIGURE D.22 – Tâche utilisée pour tester la première architecture. L'Homme et le robot doivent enlever les objets de la table et les mettre dans la boite rose.

réalisée en boucle dans un temps imparti. Le critère principal utilisé ici pour évaluer le système est le nombre de fois qu'il est capable de réaliser la tâche dans ce temps imparti. Les résultats pour 10 simulations de 30 minutes dans chaque condition peuvent être trouvés Fig. D.23(a). On peut observer une faible performance du MF seul (algorithme d'apprentissage) due à son manque de connaissances initial. La combinaison des deux experts a de bien meilleures performances bien qu'elles restent en dessous de celles d'HATP seul. En effet, la tâche étant simple à résoudre pour HATP, son plan est toujours optimal. Finalement, nous pouvons voir Fig. D.23(b) que la combinaison du MF et d'HATP permet au MF d'apprendre bien plus vite que quand il est seul.



(a) Moyenne des récompense obtenues sur 10 simulations de 30 minutes (une récompense par tâche achevée).



(b) Moyenne de l'évolution des poids de connexion du MF seul et avec HATP. Plus l'amplitude est élevée et plus le MF a appris qu'elle action exécuter.

FIGURE D.23 – Performances de l'architecture testée comparées aux experts seuls.

Les premiers résultats obtenus montrent que la combinaison d'HATP et du MF permet d'accélérer l'apprentissage du MF. Cependant, la tâche étant très simple, HATP n'a pas de difficulté quand il décide seul.

D.4.2.4 Seconde architecture : les limitations

Dans un second temps, nous avons amélioré l'architecture et l'avons testée sur une tâche plus complexe afin de démontrer l'intérêt du système.

Architecture : Comme l'un des principaux avantages du MF par rapport à HATP est son temps de calcul, nous avons modifié l'architecture comme représenté Fig. D.25. Dans cette version de l'architecture le meta-contrôleur est en amont des experts. Un expert ne sera activé uniquement que lorsque le meta-contrôleur choisira qu'il doit décider de la prochaine action.

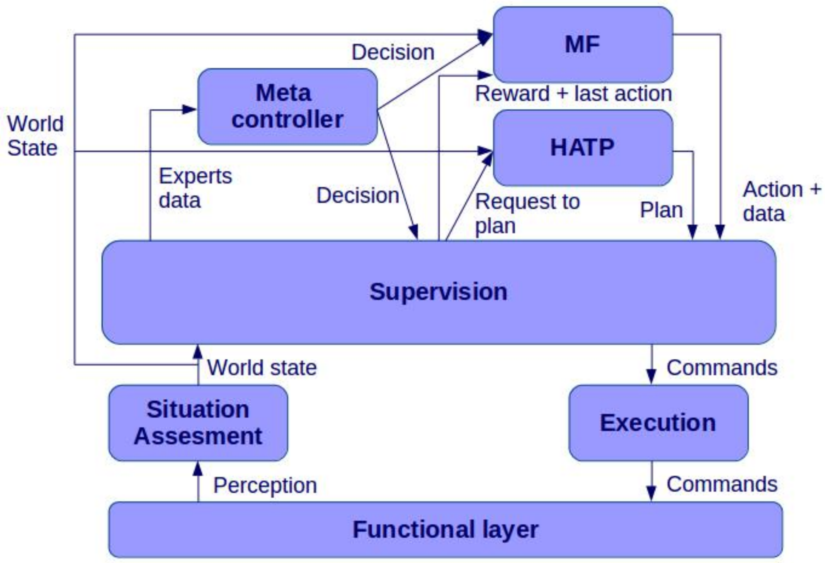


Figure D.24 – Seconde architecture développée pour combiner les deux experts.

Nous avons également introduit dans cette nouvelle architecture un nouveau critère pour la prise de décision du meta-contrôleur (précédemment aléatoire). Ce critère est basé sur le coût de chaque expert (temps à trouver une solution) et son erreur de prédiction.

Tâche : Pour augmenter la complexité de la tâche, nous avons dans un premier temps augmenté sa combinatoire. Dans la nouvelle tâche il y a maintenant 6 objets qui doivent aller dans deux boites différentes en fonction de leur couleur. Ces objets sont initialement placés de manière aléatoire sur 7 emplacements possibles sur la table au début de la tâche comme Fig. D.25. De nouvelles actions sont également possibles pour le robot, il peut maintenant replacer un objet sur un des emplacements ou naviguer à une autre position près de la table.

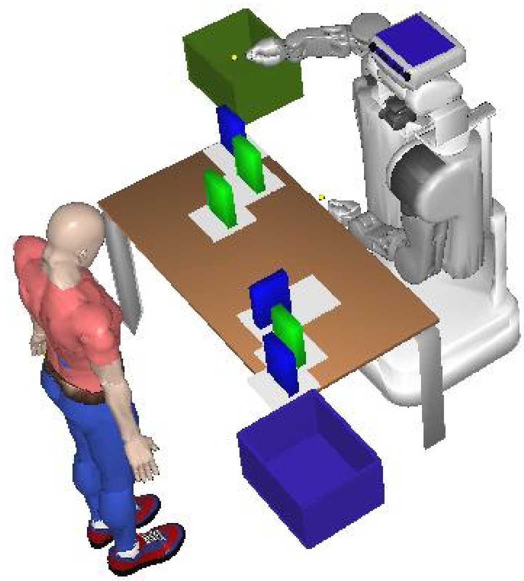
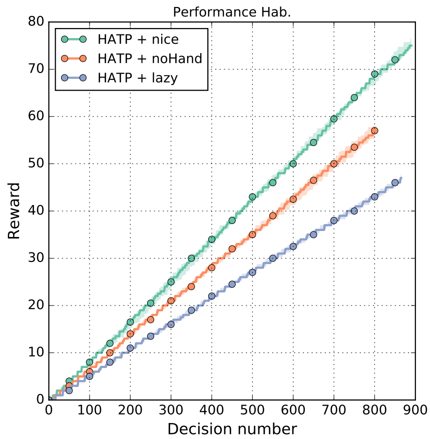


Figure D.25 – Tâche utilisée pour tester la seconde architecture. L'Homme et le robot doivent enlever les objets de la table et les mettre dans les boites de même couleur.

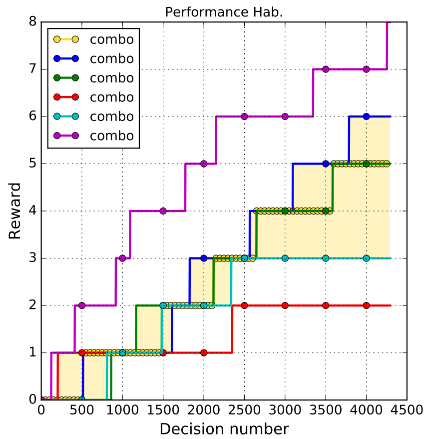
Nous avons également ajouté dans la tâche des difficultés géométriques qui ne peuvent pas être gérées par la planification (certains objets supposés accessibles ne peuvent en fait pas être attrapés par le robot). Finalement, nous avons également implémenté différents comportements pour l'homme simulé (plus ou moins coopératifs).

Résultats : Comme la tâche est plus complexe, nous avons légèrement augmenté le temps de simulation. Comme précédemment, nous avons testé le système avec chaque expert séparément puis avec la combinaison des deux. Logiquement, HATP présente de meilleurs résultats avec un Homme plus collaboratif et le MF présente de pauvres résultats seul. Cependant, nous n'avons pas réussi à avoir des résultats pour la combinaison des deux experts supérieurs à ceux d'HATP seul. Cela est dû au fait qu'avec une tâche plus complexe, l'effet d'accélération d'HATP sur l'apprentissage du MF n'est plus suffisant pour lui permettre d'apprendre une solution pour la tâche et donc lui permettre d'aider le système.

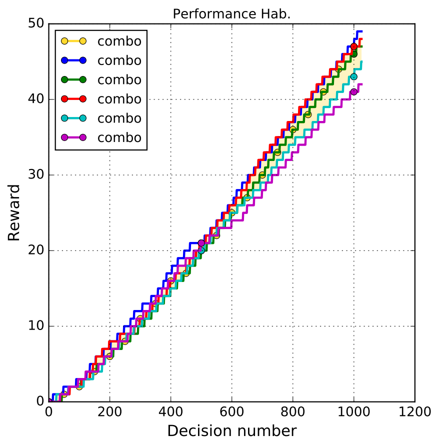
Afin d'obtenir un système plus performant, de possibles améliorations seraient de retravailler l'algorithme d'apprentissage afin de mieux l'adapter au contexte et lui permettre d'apprendre plus rapidement. Une autre amélioration possible serait de chercher un nouveau critère d'arbitrage pour le meta-contrôleur. Enfin, il serait intéressant de permettre à HATP d'obtenir un retour de ce qu'apprend le MF afin d'adapter en ligne ses modèles de planification.



(a) Moyenne des récompenses obtenues sur 10 simulations de 40 minutes pour HATP seul (une récompense par tâche achevée).



(b) Moyenne des récompenses obtenues sur 10 simulations de 40 minutes pour le MF seul (une récompense par tâche achevée).



(c) Moyenne des récompenses obtenues sur 10 simulations de 40 minutes pour la combinaison des deux experts (une récompense par tâche achevée).

FIGURE D.26 – Moyenne des récompenses obtenues dans chaque condition.

D.5 Conclusion

Plusieurs contributions sont présentées dans ce manuscrit :

— Dans un premier temps, nous avons étudié les bases de l'action conjointe entre Hommes et comment ces principes s'appliquent à l'interaction Homme-robot afin de construire un superviseur pour la décision lors de l'action conjointe Homme-robot.

— Dans un second temps nous avons étudié comment améliorer la gestion des plans partagés par le robot. Nous avons d'abord étendu l'estimation des états mentaux de l'Homme par le robot au plan partagé, puis, nous avons amélioré les algorithmes de gestion du plan pour rendre le comportement du robot plus flexible. Enfin, nous avons évalué ces deux contributions en simulation et en conditions réelles grâce à une étude utilisateur.

— Enfin, nous avons présenté deux autres contributions à l'interaction Homme-Robot. La première concerne la gestion de la tête du robot lors d'une tâche collaborative. La seconde cherche à combiner deux méthodes de prise de décision par le robot : la planification et l'apprentissage.

Bibliography

[Alami 1998] Rachid Alami, Raja Chatila, Sara Fleury, Malik Ghallab and Félix Ingrand. *An architecture for autonomy.* The International Journal of Robotics Research, vol. 17, no. 4, pages 315–337, 1998. (Cited in pages 23, 24, and 155.)

[Allen 2002] James Allen and George Ferguson. *Human-machine collaborative planning.* In Third International NASA Workshop on Planning and Scheduling for Space, 2002. (Cited in pages 20, 63, and 154.)

[Anderson 2004] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere and Yulin Qin. *An integrated theory of the mind.* Psychological review, vol. 111, no. 4, page 1036, 2004. (Cited in page 45.)

[Argyle 1972] Michael Argyle and Roger Ingham. *Gaze, mutual gaze, and proximity.* Semiotica, vol. 6, no. 1, pages 32–49, 1972. (Cited in page 100.)

[Argyle 1976] Michael Argyle and Mark Cook. *Gaze and mutual gaze.* 1976. (Cited in pages 100 and 174.)

[Baker 2014] Chris L Baker and Joshua B Tenenbaum. *Modeling human plan recognition using bayesian theory of mind.* Plan, activity, and intent recognition: Theory and practice, pages 177–204, 2014. (Cited in pages 19, 20, 46, 154, and 160.)

[Balleine 2010] Bernard W Balleine and John P O'doherty. *Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action.* Neuropsychopharmacology, vol. 35, no. 1, pages 48–69, 2010. (Cited in page 118.)

[Baraglia 2016] Jimmy Baraglia, Maya Cakmak, Yukie Nagai, Rajesh Rao and Minoru Asada. *Initiative in robot assistance during collaborative task execution.* In HRI, ACM/IEEE, 2016. (Cited in page 62.)

[Baron-Cohen 1985] Simon Baron-Cohen, Alan M Leslie and Uta Frith. *Does the autistic child have a "theory of mind"?* Cognition, vol. 21, no. 1, pages 37–46, 1985. (Cited in pages 44 and 159.)

[Baron-Cohen 1995] Simon Baron-Cohen. *The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology.* In Portions of this paper were presented at the Society for Research in Child Development Conference, New Orleans, Mar 1993; the British Psychological Society, Welsh Branch," Faces" Conference, U Wales Coll of Cardiff, Sep 1993; and the British Society for the Philosophy of Science" Roots of Joint Reference" Conference, U Bristol, Nov 1993. Lawrence Erlbaum Associates, Inc, 1995. (Cited in page 100.)

[Baron-Cohen 1997] Simon Baron-Cohen. Mindblindness: An essay on autism and theory of mind. MIT press, 1997. (Cited in pages 45 and 160.)

[Bartneck 2009] Christoph Bartneck, Dana Kulić, Elizabeth Croft and Susana Zoghbi. *Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots.* International journal of social robotics, vol. 1, no. 1, pages 71–81, 2009. (Cited in pages 86, 87, and 171.)

[Baxter 2013] Paul E Baxter, Joachim de Greeff and Tony Belpaeme. *Cognitive architecture for human–robot interaction: towards behavioural alignment.* Biologically Inspired Cognitive Architectures, vol. 6, pages 30–39, 2013. (Cited in page 26.)

[Becchio 2010] Cristina Becchio, Luisa Sartori and Umberto Castiello. *Toward you: The social side of actions.* Current Directions in Psychological Science, vol. 19, no. 3, pages 183–188, 2010. (Cited in pages 100, 102, and 174.)

[Beetz 2010] Michael Beetz, Lorenz Mösenlechner and Moritz Tenorth. *CRAM—A Cognitive Robot Abstract Machine for everyday manipulation in human environments.* In Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, pages 1012–1017. IEEE, 2010. (Cited in page 26.)

[Berlin 2006] Matt Berlin, Jesse Gray, Andrea Lockerd Thomaz and Cynthia Breazeal. *Perspective taking: An organizing principle for learning in human-robot interaction.* In AAAI, volume 2, pages 1444–1450, 2006. (Cited in pages 45 and 160.)

[Botelho 1999] Sylvia C Botelho and Rachid Alami. *M+: a scheme for multi-robot cooperation through negotiated task allocation and achievement.* In Robotics and Automation, IEEE, 1999. (Cited in page 63.)

[Boucher 2010] Jean-David Boucher, Jocelyne Ventre-Dominey, Peter Ford Dominey, Sacha Fagel and Gerard Bailly. *Facilitative effects of communicative gaze and speech in human-robot cooperation.* In Proceedings of the 3rd international workshop on Affective interaction in natural environments, pages 71–74. ACM, 2010. (Cited in pages 21, 101, 102, 154, and 174.)

[Boucher 2012] Jean-David Boucher, Ugo Pattacini, Amelie Lelong, Gerard Bailly, Frederic Elisei, Sascha Fagel, Peter Ford Dominey and Jocelyne Ventre-Dominey. *I reach faster when I see you look: gaze effects in human–human and human–robot face-to-face cooperation.* Frontiers in neurorobotics, vol. 6, 2012. (Cited in pages 101 and 174.)

[Brafman 2002] Ronen I Brafman and Moshe Tennenholtz. *R-max-a general polynomial time algorithm for near-optimal reinforcement learning.* Journal of Machine Learning Research, vol. 3, no. Oct, pages 213–231, 2002. (Cited in page 121.)

[Bratman 1989] Michael E Bratman. *Intention and personal policies.* Philosophical perspectives, vol. 3, pages 443–469, 1989. (Cited in pages 12 and 152.)

[Bratman 1993] Michael E Bratman. *Shared intention.* Ethics, vol. 104, no. 1, pages 97–113, 1993. (Cited in pages 13, 51, and 152.)

[Breazeal 2005] Cynthia Breazeal, Cory D Kidd, Andrea Lockerd Thomaz, Guy Hoffman and Matt Berlin. *Effects of nonverbal communication on efficiency and robustness in human-robot teamwork*. In Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on, pages 708–713. IEEE, 2005. (Cited in pages 21, 101, and 154.)

[Breazeal 2006] Cynthia Breazeal, Matt Berlin, Andrew Brooks, Jesse Gray and Andrea L Thomaz. *Using perspective taking to learn from ambiguous demonstrations*. Robotics and autonomous systems, vol. 54, no. 5, pages 385–393, 2006. (Cited in pages 20, 46, 154, and 160.)

[Breazeal 2009] Cynthia Breazeal, Jesse Gray and Matt Berlin. *An embodied cognition approach to mindreading skills for socially intelligent robots*. The International Journal of Robotics Research, vol. 28, no. 5, pages 656–680, 2009. (Cited in pages 19 and 154.)

[Broekens 2013] Joost Broekens and Willem-Paul Brinkman. *AffectButton: A method for reliable and valid affective self-report*. International Journal of Human-Computer Studies, vol. 71, no. 6, pages 641–667, 2013. (Cited in pages 87, 88, and 171.)

[Brooke 1986] John Brooke. *System usability scale (SUS): a quick-and-dirty method of system evaluation user information*. Reading, UK: Digital Equipment Co Ltd, 1986. (Cited in page 86.)

[Buekens 2001] FAI Buekens, X Vanmechelen and K Maessen. *Indexicaliteit en dynamische intenties*. 2001. (Cited in page 22.)

[Bui 2003] Hung Hai Bui. *A general model for online probabilistic plan recognition*. In IJCAI, volume 3, pages 1309–1315, 2003. (Cited in pages 19 and 154.)

[Butterworth 1991] George Butterworth. *The ontogeny and phylogeny of joint visual attention*. 1991. (Cited in page 100.)

[Cakmak 2011] Maya Cakmak, Siddhartha S Srinivasa, Min Kyung Lee, Jodi Forlizzi and Sara Kiesler. *Human preferences for robot-human hand-over configurations*. In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, pages 1986–1993. IEEE, 2011. (Cited in pages 21 and 154.)

[Castiello 2003] Umberto Castiello. *Understanding other people's actions: intention and attention*. Journal of Experimental Psychology: Human Perception and Performance, vol. 29, no. 2, page 416, 2003. (Cited in pages 100 and 174.)

[Chakraborti 2015] Tathagata Chakraborti, Gordon Briggs, Kartik Talamadupula, Matthias Scheutz, David Smith and Subbarao Kambhampati. *Planning for serendipity-altruism in human-robot cohabitation*. In ICAPS Workshop on Planning and Robotics (PlanRob), 2015. (Cited in pages 19 and 62.)

[Chakraborti 2016] Tathagata Chakraborti, Yu Zhang, David E Smith and Subbarao Kambhampati. *Planning with resource conflicts in human-robot cohabitation*. In AAMAS, 2016. (Cited in pages 19 and 62.)

[Charisi 2017] Vicky Charisi, Bram Ridder, Jaebok Kim and Vanessa Evers. *Planning Based System for Child-Robot Interaction in Dynamic Play Environments.* In Robotics: Science and Systems, 2017. (Cited in page 19.)

[Chien 2000] Steve A Chien, Russell Knight, Andre Stechert, Rob Sherwood and Gregg Rabideau. *Using Iterative Repair to Improve the Responsiveness of Planning and Scheduling.* In AIPS, 2000. (Cited in page 63.)

[Cirillo 2010] Marcello Cirillo, Lars Karlsson and Alessandro Saffiotti. *Human-aware task planning: an application to mobile robots.* IST, 2010. (Cited in pages 19, 62, 119, and 154.)

[Clark 1996] Herbert H Clark. *Using language. 1996.* Cambridge University Press: Cambridge, vol. 952, pages 274–296, 1996. (Cited in pages 18, 21, and 154.)

[Clodic 2009] Aurélie Clodic, Hung Cao, Samir Alili, Vincent Montreuil, Rachid Alami and Raja Chatila. *Shary: a supervision system adapted to human-robot interaction.* In Experimental Robotics, pages 229–238. Springer, 2009. (Cited in pages 21, 31, 63, and 154.)

[Clodic 2014] Aurélie Clodic, Rachid Alami and Raja Chatila. *Key Elements for Human-Robot Joint Action.* Frontiers in AI and Applications, 2014. (Cited in pages 72 and 165.)

[Cohen 1991] Philip R Cohen and Hector J Levesque. *Teamwork.* Nous, vol. 25, no. 4, pages 487–512, 1991. (Cited in pages 12, 51, and 160.)

[Cook 1977] Mark Cook. *Gaze and Mutual Gaze in Social Encounters: How long—and when—we look others" in the eye" is one of the main signals in nonverbal communication.* American Scientist, vol. 65, no. 3, pages 328–333, 1977. (Cited in page 100.)

[Coradeschi 2003] Silvia Coradeschi and Alessandro Saffiotti. *An introduction to the anchoring problem.* Robotics and Autonomous Systems, vol. 43, no. 2, pages 85–96, 2003. (Cited in pages 20 and 154.)

[Cronbach 1951] Lee J Cronbach. *Coefficient alpha and the internal structure of tests.* psychometrika, vol. 16, no. 3, pages 297–334, 1951. (Cited in page 91.)

[Dautenhahn 2007] Kerstin Dautenhahn. *Socially intelligent robots: dimensions of human–robot interaction.* Philosophical Transactions of the Royal Society of London B: Biological Sciences, vol. 362, no. 1480, pages 679–704, 2007. (Cited in pages 1 and 151.)

[Davis 1989] Fred D Davis. *Perceived usefulness, perceived ease of use, and user acceptance of information technology.* MIS quarterly, pages 319–340, 1989. (Cited in pages 85, 86, and 171.)

[Daw 2005] Nathaniel D Daw, Yael Niv and Peter Dayan. *Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control.* Nature neuroscience, vol. 8, no. 12, pages 1704–1711, 2005. (Cited in page 119.)

[De Weerd 2014] Harmen De Weerd, Rineke Verbrugge and Bart Verheij. *Theory of Mind in the Mod Game: An Agent-Based Model of Strategic Reasoning.* In ECSI, pages 128–136, 2014. (Cited in page 44.)

[Dehais 2011] Frédéric Dehais, Emrah Akin Sisbot, Rachid Alami and Mickaël Causse. *Physiological and subjective evaluation of a human–robot object hand-over task.* Applied ergonomics, vol. 42, no. 6, pages 785–791, 2011. (Cited in pages 21 and 154.)

[Demiris 2006] Yiannis Demiris and Bassam Khadhouri. *Hierarchical attentive multiple models for execution and recognition of actions.* Robotics and autonomous systems, vol. 54, no. 5, pages 361–369, 2006. (Cited in page 26.)

[DePaulo 1992] Bella M DePaulo. *Nonverbal behavior and self-presentation.* Psychological bulletin, vol. 111, no. 2, page 203, 1992. (Cited in pages 99, 100, and 174.)

[Devin 2016a] Sandra Devin and Rachid Alami. *An implemented theory of mind to improve human-robot shared plans execution.* In Human-Robot Interaction (HRI), 2016 11th ACM/IEEE International Conference on, pages 319–326. IEEE, 2016. (Cited in page 44.)

[Devin 2016b] Sandra Devin, Grégoire Milliez, Michelangelo Fiore, Aurélie Clodic and Rachid Alami. *Some essential skills and their combination in an architecture for a cognitive and interactive robot.* arXiv preprint arXiv:1603.00583, 2016. (Cited in pages 24 and 156.)

[Devin 2017] Sandra Devin, Aurélie Clodic and Rachid Alami. *About Decisions During Human-Robot Shared Plan Achievement: Who Should Act and How?* In ICSR'17. to appear, 2017. (Cited in page 62.)

[Dickinson 1985] Anthony Dickinson. *Actions and habits: the development of behavioural autonomy.* Philosophical Transactions of the Royal Society of London B: Biological Sciences, vol. 308, no. 1135, pages 67–78, 1985. (Cited in pages 118 and 179.)

[Dillon 2001] Andrew Dillon. *User acceptance of information technology.* In Encyclopedia of human factors and ergonomics. London: Taylor and Francis, 2001. (Cited in page 85.)

[Dolan 2013] Ray J Dolan and Peter Dayan. *Goals and habits in the brain.* Neuron, vol. 80, no. 2, pages 312–325, 2013. (Cited in page 118.)

[Dollé 2010] Laurent Dollé, Denis Sheynikhovich, Benoît Girard, Ricardo Chavarriaga and Agnès Guillot. *Path planning versus cue responding: a bio-inspired model of switching between navigation strategies.* Biological cybernetics, vol. 103, no. 4, pages 299–317, 2010. (Cited in page 119.)

[Dragan 2015] Anca D Dragan, Shira Bauman, Jodi Forlizzi and Siddhartha S Srinivasa. *Effects of robot motion on human-robot collaboration.* In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, pages 51–58. ACM, 2015. (Cited in page 86.)

[Ekman 1969] Paul Ekman and Wallace V Friesen. *The repertoire of nonverbal behavior: Categories, origins, usage, and coding.* semiotica, vol. 1, no. 1, pages 49–98, 1969. (Cited in pages 99, 100, and 174.)

[Emery 2000] Nathan J Emery. *The eyes have it: the neuroethology, function and evolution of social gaze.* Neuroscience & Biobehavioral Reviews, vol. 24, no. 6, pages 581–604, 2000. (Cited in page 100.)

[Erol 1994] Kutluhan Erol, James Hendler and Dana S Nau. *HTN planning: Complexity and expressivity.* In AAAI, volume 94, pages 1123–1128, 1994. (Cited in page 120.)

[Ferreira 2015] Emmanuel Ferreira, Grégoire Milliez, Fabrice Lefèvre and Rachid Alami. *Users' belief awareness in reinforcement learning-based situated human–robot dialogue management.* In Natural Language Dialog Systems and Intelligent Assistants, pages 73–86. Springer, 2015. (Cited in pages 20, 21, and 154.)

[Fiore 2014] Michelangelo Fiore, Aurélie Clodic and Rachid Alami. *On planning and task achievement modalities for human-robot collaboration.* In ISER, 2014. (Cited in pages 63 and 164.)

[Fiore 2016] Michelangelo Fiore, Aurélie Clodic and Rachid Alami. *On planning and task achievement modalities for human-robot collaboration.* In Experimental Robotics, pages 293–306. Springer, 2016. (Cited in page 31.)

[Fischer 2016] Kerstin Fischer, Lars C Jensen, Stefan-Daniel Suvei and Leon Bodenhagen. *Between legibility and contact: The role of gaze in robot approach.* In Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on, pages 646–651. IEEE, 2016. (Cited in pages 86 and 101.)

[Flavell 1977] John H Flavell. *The development of knowledge about visual perception.* In Nebraska symposium on motivation. University of Nebraska Press, 1977. (Cited in pages 44 and 159.)

[Flavell 1992] John H Flavell, H Beilin and P Pufall. *Perspectives on perspective-taking.* Piaget's theory: Prospects and possibilities, pages 107–139, 1992. (Cited in pages 44 and 159.)

[Frick 2014] Andrea Frick, Wenke Möhring and Nora S Newcombe. *Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds.* Frontiers in psychology, vol. 5, page 386, 2014. (Cited in page 44.)

[Furlanetto 2013] Tiziano Furlanetto, Andrea Cavallo, Valeria Manera, Barbara Tversky and Cristina Becchio. *Through your eyes: incongruence of gaze and action increases spontaneous perspective taking.* Frontiers in human neuroscience, vol. 7, 2013. (Cited in pages 101 and 174.)

[Georgeff 1987] Michael P Georgeff and Amy L Lansky. *Reactive reasoning and planning.* In AAAI, volume 87, pages 677–682, 1987. (Cited in page 19.)

[Gerkey 2002] Brian P Gerkey and Maja J Mataric. *Sold!: Auction methods for multirobot coordination.* IEEE Trans. on Robotics and Automation, 2002. (Cited in page 63.)

[Gerkey 2004] Brian P Gerkey and Maja J Matarić. *A formal analysis and taxonomy of task allocation in multi-robot systems.* The IJRResearch, 2004. (Cited in page 63.)

[Ghallab 1994] Malik Ghallab and Hervé Laruelle. *Representation and Control in IxTeT, a Temporal Planner.* In AIPS, volume 1994, pages 61–67, 1994. (Cited in pages 19 and 154.)

[Gharbi 2015] Mamoun Gharbi, Pierre-Vincent Paubel, Aurélie Clodic, Ophélie Carreras, Rachid Alami and Jean-Marie Cellier. *Toward a better understanding of the communication cues involved in a human-robot object transfer.* In Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on, pages 319–324. IEEE, 2015. (Cited in pages 101 and 102.)

[Gibson 1977] James J Gibson. *Perceiving, acting, and knowing: Toward an ecological psychology.* The Theory of Affordances, pages 67–82, 1977. (Cited in pages 17 and 153.)

[Goffman 1979] Erving Goffman. *Footing.* Semiotica, vol. 25, no. 1-2, pages 1–30, 1979. (Cited in page 100.)

[Gombolay 2015] Matthew C Gombolay, Reymundo A Gutierrez, Shanelle G Clarke, Giancarlo F Sturla and Julie A Shah. *Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams.* Autonomous Robots, 2015. (Cited in pages 62 and 86.)

[Görür 2017] O Can Görür, Benjamin Rosman, Guy Hoffman and Şahin Albayrak. *Toward Integrating Theory of Mind into Adaptive Decision-Making of Social Robots to Understand Human Intention.* 2017. (Cited in page 46.)

[Gray 2014] Jesse Gray and Cynthia Breazeal. *Manipulating mental states through physical action.* International Journal of Social Robotics, vol. 6, no. 3, pages 315–327, 2014. (Cited in pages 20 and 46.)

[Grosz 1988] Barbara J Grosz and Candace L Sidner. *Plans for discourse.* Technical Report, DTIC Document, 1988. (Cited in pages 14 and 152.)

[Grosz 1999] Barbara J Grosz and Sarit Kraus. *The evolution of SharedPlans.* In Foundations of rational agency, pages 227–262. Springer, 1999. (Cited in page 14.)

[Häring 2012] Markus Häring, Jessica Eichberg and Elisabeth André. *Studies on grounding with gaze and pointing gestures in human-robot-interaction.* Social Robotics, pages 378–387, 2012. (Cited in pages 101 and 174.)

[Hart 2014] Justin W Hart, Brian Gleeson, Matthew Pan, AJung Moon, Karon MacLean and Elizabeth Croft. *Gesture, gaze, touch, and hesitation: Timing cues for collaborative work.* In HRI Workshop on Timing in Human-Robot Interaction, Bielefeld, Germany, 2014. (Cited in pages 21 and 154.)

[Hassenzahl 2003] Marc Hassenzahl. *The thing and I: understanding the relationship between user and product.* In Funology, pages 31–42. Springer, 2003. (Cited in pages 85 and 171.)

[Hassenzahl 2006] Marc Hassenzahl and Noam Tractinsky. *User experience-a research agenda.* Behaviour & information technology, vol. 25, no. 2, pages 91–97, 2006. (Cited in page 84.)

[Hayes 1994] Gillian M Hayes and John Demiris. A robot controller using learning by imitation. University of Edinburgh, Department of Artificial Intelligence, 1994. (Cited in page 119.)

[Heerink 2009] Marcel Heerink, Ben Krose, Vanessa Evers and Bob Wielinga. *Measuring acceptance of an assistive social robot: a suggested toolkit.* In Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on, pages 528–533. IEEE, 2009. (Cited in page 86.)

[Heerink 2010] Marcel Heerink, Ben Kröse, Vanessa Evers and Bob Wielinga. *Relating conversational expressiveness to social presence and acceptance of an assistive social robot.* Virtual reality, vol. 14, no. 1, pages 77–84, 2010. (Cited in page 86.)

[Hemsley 1978] Gordon D Hemsley and Anthony N Doob. *The effect of looking behavior on perceptions of a communicator's credibility.* Journal of Applied Social Psychology, vol. 8, no. 2, pages 136–142, 1978. (Cited in page 101.)

[Hiatt 2004] Laura M Hiatt, J Gregory Trafton, Anthony M Harrison and Alan C Schultz. *A Cognitive Model for Spatial Perspective taking.* In ICCM, pages 354–355, 2004. (Cited in page 45.)

[Hiatt 2010] Laura M Hiatt and J Gregory Trafton. *A cognitive model of theory of mind.* In Proceedings of the 10th International Conference on Cognitive Modeling, pages 91–96. Citeseer, 2010. (Cited in pages 45 and 160.)

[Hiatt 2011] Laura M Hiatt, Anthony M Harrison and J Gregory Trafton. *Accommodating human variability in human-robot teams through theory of mind.* In IJCAI Proceedings-International Joint Conference on Artificial Intelligence, volume 22, page 2066, 2011. (Cited in page 45.)

[Hiatt 2015] Laura M Hiatt and J Gregory Trafton. *Understanding second-order theory of mind.* In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, pages 167–168. ACM, 2015. (Cited in page 45.)

[Hoffman 2007] Guy Hoffman and Cynthia Breazeal. *Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team.* In HRI, ACM/IEEE, 2007. (Cited in page 63.)

[Hoffman 2013] Guy Hoffman. *Evaluating fluency in human-robot collaboration.* In International conference on human-robot interaction (HRI), workshop on human robot collaboration, volume 381, pages 1–8, 2013. (Cited in page 86.)

[Imai 2002] Michita Imai, Takayuki Kanda, Testuo Ono, Hiroshi Ishiguro and Kenji Mase. *Robot mediated round table: Analysis of the effect of robot's gaze.* In Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on, pages 411–416. IEEE, 2002. (Cited in pages 100, 101, and 174.)

[Ingrand 2014] Félix Ingrand and Malik Ghallab. *Deliberation for autonomous robots: A survey.* Artificial Intelligence, 2014. (Cited in pages 119 and 179.)

[Johnson 2005a] Matthew Johnson and Yiannis Demiris. *Perceptual perspective taking and action recognition.* International Journal of Advanced Robotic Systems, vol. 2, no. 4, page 32, 2005. (Cited in pages 45 and 160.)

[Johnson 2005b] MR Johnson and YK Demiris. *Perspective taking through simulation.* 2005. (Cited in page 45.)

[Joyeux 2009] Sylvain Joyeux, Rachid Alami, Simon Lacroix and Roland Philippsen. *A plan manager for multi-robot systems.* IJRR, 2009. (Cited in page 63.)

[Kaplan 2002] Frederic Kaplan, Pierre-Yves Oudeyer, Enikö Kubinyi and Adám Miklósi. *Robotic clicker training.* Robotics and Autonomous Systems, vol. 38, no. 3, pages 197–206, 2002. (Cited in page 119.)

[Karpas 2015] Erez Karpas, Steven James Levine, Peng Yu and Brian C Williams. *Robust Execution of Plans for Human-Robot Teams.* In ICAPS, pages 342–346, 2015. (Cited in pages 21, 63, and 154.)

[Kendon 1967] Adam Kendon. *Some functions of gaze-direction in social interaction.* Acta psychologica, vol. 26, pages 22–63, 1967. (Cited in pages 100 and 174.)

[Keramati 2011] Mehdi Keramati, Amir Dezfouli and Payam Piray. *Speed/accuracy trade-off between the habitual and the goal-directed processes.* PLoS Comput Biol, vol. 7, no. 5, page e1002055, 2011. (Cited in page 119.)

[Khamassi 2016] Mehdi Khamassi, Benoît Girard, Aurélie Clodic, Sandra Devin, Erwan Renaudo, Elisabeth Pacherie, Rachid Alami and Raja Chatila. *Integration of Action, Joint Action and Learning in Robot Cognitive Architectures.* Intellectica-La revue de l'Association pour la Recherche sur les sciences de la Cognition (ARCo), vol. 2016, no. 65, pages 169–203, 2016. (Cited in page 118.)

[Khambhaita 2016] Harmish Khambhaita, Jorge Rios-Martinez and Rachid Alami. *Head-Body Motion Coordination for Human Aware Robot Navigation.* In Proc. International Workshop on Human-Friendlly Robotics (HFR), 2016. (Cited in page 101.)

[Kim 2015] Hyun K Kim, Sung H Han, Jaehyun Park and Wonkyu Park. *How user experience changes over time: A case study of social network services.* Human Factors and Ergonomics in Manufacturing & Service Industries, vol. 25, no. 6, pages 659–673, 2015. (Cited in page 96.)

[Kleinke 1986] Chris L Kleinke. *Gaze and eye contact: a research review.* Psychological bulletin, vol. 100, no. 1, page 78, 1986. (Cited in page 101.)

[Knoblich 2011] Günther Knoblich, Stephen Butterfill and Natalie Sebanz. *3 Psychological research on joint action: theory and data.* Psychology of Learning and Motivation-Advances in Research and Theory, vol. 54, page 59, 2011. (Cited in pages 17 and 153.)

[Knox 2012] W Bradley Knox and Peter Stone. *Reinforcement learning from simultaneous human and MDP reward.* In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1, pages 475–482. International Foundation for Autonomous Agents and Multiagent Systems, 2012. (Cited in page 119.)

[Kober 2011] Jens Kober and Jan Peters. *Learning elementary movements jointly with a higher level task.* In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, pages 338–343. IEEE, 2011. (Cited in pages 119 and 179.)

[Kruse 2013] Thibault Kruse, Amit Kumar Pandey, Rachid Alami and Alexandra Kirsch. *Human-aware robot navigation: A survey.* Robotics and Autonomous Systems, vol. 61, no. 12, pages 1726–1743, 2013. (Cited in pages 21 and 155.)

[LaBarre 1947] Weston LaBarre. *The cultural basis of emotions and gestures.* Journal of personality, vol. 16, no. 1, pages 49–68, 1947. (Cited in pages 100 and 174.)

[Lallée 2013] Stéphane Lallée, Katharina Hamann, Jasmin Steinwender, Felix Warneken, Uriel Martienz, Hector Barron-Gonzales, Ugo Pattacini, Ilaria Gori, Marc Petit, Giorgio Metta*et al.* *Cooperative human robot interaction systems: IV. Communication of shared plans with Naïve humans using gaze and speech.* In IROS, IEEE/RSJ, 2013. (Cited in pages 62, 101, 103, and 174.)

[Lallemand 2015] Carine Lallemand, Vincent Koenig, Guillaume Gronier and Romain Martin. *Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs.* Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, vol. 65, no. 5, pages 239–252, 2015. (Cited in pages 87 and 171.)

[Lallement 2014] Raphaël Lallement, Lavindra de Silva and Rachid Alami. *HATP: An HTN Planner for Robotics.* CoRR, 2014. (Cited in pages 19, 30, 62, 64, 68, 119, 154, 163, and 179.)

[Lemai 2004] Solange Lemai and Félix Ingrand. *Interleaving temporal planning and execution in robotics domains.* In AAAI, volume 4, pages 617–622, 2004. (Cited in pages 19 and 154.)

[Lemaignan 2012] Séverin Lemaignan, Raquel Ros, E Akin Sisbot, Rachid Alami and Michael Beetz. *Grounding the interaction: Anchoring situated dis-*

course in everyday human-robot interaction. International Journal of Social Robotics, vol. 4, no. 2, pages 181–199, 2012. (Cited in pages 20 and 154.)

[Lesaint 2014] Florian Lesaint, Olivier Sigaud, Shelly B Flagel, Terry E Robinson and Mehdi Khamassi. *Modelling individual differences in the form of pavlovian conditioned approach responses: a dual learning systems approach with factored representations.* PLoS Comput Biol, vol. 10, no. 2, page e1003466, 2014. (Cited in pages 119 and 179.)

[Leslie 1984] Alan M Leslie. *Spatiotemporal continuity and the perception of causality in infants.* Perception, vol. 13, no. 3, pages 287–305, 1984. (Cited in pages 45 and 160.)

[Levine 2014] Steven James Levine and Brian Charles Williams. *Concurrent Plan Recognition and Execution for Human-Robot Teams.* In ICAPS, 2014. (Cited in page 63.)

[Lucignano 2013] Lorenzo Lucignano, Francesco Cutugno, Silvia Rossi and Alberto Finzi. *A dialogue system for multimodal human-robot interaction.* In Proceedings of the 15th ACM on International conference on multimodal interaction, pages 197–204. ACM, 2013. (Cited in pages 21 and 154.)

[Mahlke 2008] Sascha Mahlke. *User experience of interaction with technical systems.* 2008. (Cited in pages 85 and 171.)

[Mainprice 2012] Jim Mainprice, Mamoun Gharbi, Thierry Siméon and Rachid Alami. *Sharing effort in planning human-robot handover tasks.* In RO-MAN, 2012 IEEE, pages 764–770. IEEE, 2012. (Cited in pages 21 and 154.)

[Martins 2010] Murilo Fernandes Martins and Yiannis Demiris. *Learning multirobot joint action plans from simultaneous task execution demonstrations.* In Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1, pages 931–938. International Foundation for Autonomous Agents and Multiagent Systems, 2010. (Cited in pages 119 and 179.)

[Mason 2005] Andrea H Mason and Christine L MacKenzie. *Grip forces when passing an object to a partner.* Experimental brain research, vol. 163, no. 2, pages 173–187, 2005. (Cited in page 21.)

[Mavridis 2005] Nikolaos Mavridis and Deb Roy. *Grounded situation models for robots: Bridging language, perception, and action.* In AAAI-05 workshop on modular construction of human-like intelligence, 2005. (Cited in pages 20 and 154.)

[Mehrabian 1969] Albert Mehrabian. *Significance of posture and position in the communication of attitude and status relationships.* Psychological Bulletin, vol. 71, no. 5, page 359, 1969. (Cited in pages 100 and 174.)

[Milliez 2014] Grégoire Milliez, Matthieu Warnier, Aurélie Clodic and Rachid Alami. *A framework for endowing an interactive robot with reasoning capabilities about perspective-taking and belief management.* In Robot and

Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on, pages 1103–1109. IEEE, 2014. (Cited in pages 20, 45, 50, 154, and 160.)

[Milliez 2016a] Grégoire Milliez. *Raisonnement sur le contexte et les croyances pour l'interaction homme-robot.* PhD thesis, INP-T, LAAS-CNRS, 2016. (Cited in page 29.)

[Milliez 2016b] Grégoire Milliez, Raphaël Lallement, Michelangelo Fiore and Rachid Alami. *Using human knowledge awareness to adapt collaborative plan generation, explanation and monitoring.* In HRI ACM/IEEE, 2016. (Cited in pages 20, 47, 63, and 154.)

[Mohseni-Kabir 2015] A. Mohseni-Kabir, C. Rich, S. Chernova, C. L. Sidner and D. Miller. *Interactive Hierarchical Task Learning from a Single Demonstration.* In HRI, ACM/IEEE, 2015. (Cited in pages 20, 63, and 154.)

[Molineaux 2010] Matt Molineaux, Matthew Klenk and David W Aha. *Goal-driven autonomy in a Navy strategy simulation.* Technical Report, DTIC Document, 2010. (Cited in pages 19 and 63.)

[Moon 2014] AJung Moon, Daniel M Troniak, Brian Gleeson, Matthew KXJ Pan, Minhua Zheng, Benjamin A Blumer, Karon MacLean and Elizabeth A Croft. *Meet me where i'm gazing: how shared attention gaze affects human-robot handover timing.* In Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction, pages 334–341. ACM, 2014. (Cited in pages 101 and 102.)

[Morimoto 2001] Jun Morimoto and Kenji Doya. *Acquisition of stand-up behavior by a real robot using hierarchical reinforcement learning.* Robotics and Autonomous Systems, vol. 36, no. 1, pages 37–51, 2001. (Cited in page 119.)

[Mutlu 2009a] Bilge Mutlu. *Designing gaze behavior for humanlike robots.* PhD thesis, Carnegie Mellon University, 2009. (Cited in pages 100 and 174.)

[Mutlu 2009b] Bilge Mutlu, Toshiyuki Shiwa, Takayuki Kanda, Hiroshi Ishiguro and Norihiro Hagita. *Footing in human-robot conversations: how robots might shape participant roles using gaze cues.* In Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, pages 61–68. ACM, 2009. (Cited in pages 21, 101, 102, 154, and 174.)

[Nagai 2015] Takayuki Nagai, Kasumi Abe, Tomoaki Nakamura, Natsuki Oka and Takashi Omori. *Probabilistic modeling of mental models of others.* In Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on, pages 89–94. IEEE, 2015. (Cited in pages 20, 46, 154, and 160.)

[Neider 2010] Mark B Neider, Xin Chen, Christopher A Dickinson, Susan E Brennan and Gregory J Zelinsky. *Coordinating spatial referencing using shared gaze.* Psychonomic bulletin & review, vol. 17, no. 5, pages 718–724, 2010. (Cited in page 101.)

[Nicolescu 2003] Monica N Nicolescu and Maja J Mataric. *Natural methods for robot task learning: Instructive demonstrations, generalization and practice*. In Proceedings of the second international joint conference on Autonomous agents and multiagent systems, pages 241–248. ACM, 2003. (Cited in page 119.)

[Nikolaidis 2013] Stefanos Nikolaidis and Julie Shah. *Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy*. In Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction, pages 33–40. IEEE Press, 2013. (Cited in page 119.)

[Obhi 2011] Sukhvinder S Obhi and Natalie Sebanz. *Moving together: toward understanding the mechanisms of joint action*, 2011. (Cited in pages 16 and 152.)

[Pacherie 2008] Elisabeth Pacherie. *The phenomenology of action: A conceptual framework*. Cognition, vol. 107, no. 1, pages 179–217, 2008. (Cited in pages 22 and 23.)

[Pacherie 2011] Elisabeth Pacherie. *The Phenomenology of Joint Action: Self-Agency versus Joint Agency*. Joint attention: New developments in psychology, philosophy of mind, and social neuroscience, page 343, 2011. (Cited in pages 16, 22, 152, and 155.)

[Petit 2013] Marc Petit, Stéphane Lallée, J-D Boucher, Grégoire Pointeau, Pierrick Cheminade, Dimitri Ognibene, Eris Chinellato, Ugo Pattacini, Ilaria Gori, Uriel Martinez-Hernandez*et al.* *The coordinating role of language in real-time multimodal learning of cooperative tasks*. Autonomous Mental Development, IEEE, 2013. (Cited in pages 20, 63, and 154.)

[Pezzulo 2013] Giovanni Pezzulo, Francesco Rigoli and Fabian Chersi. *The mixed instrumental controller: using value of information to combine habitual choice and mental simulation*. Frontiers in psychology, vol. 4, page 92, 2013. (Cited in pages 119 and 179.)

[Pierno 2006] Andrea C Pierno, Cristina Becchio, Matthew B Wall, Andrew T Smith, Luca Turella and Umberto Castiello. *When gaze turns into grasp*. Journal of Cognitive Neuroscience, vol. 18, no. 12, pages 2130–2137, 2006. (Cited in pages 100 and 174.)

[Pointeau 2014] Grégoire Pointeau, Maxime Petit and Peter Ford Dominey. *Successive developmental levels of autobiographical memory for learning through social interaction*. IEEE Transactions on Autonomous Mental Development, vol. 6, no. 3, pages 200–212, 2014. (Cited in pages 20 and 154.)

[Premack 1978] David Premack and Guy Woodruff. *Does the chimpanzee have a theory of mind?* Behavioral and brain sciences, vol. 1, no. 04, pages 515–526, 1978. (Cited in pages 44 and 159.)

[Rabideau 1999] Gregg Rabideau, Russell Knight, Steve Chien, Alex Fukunaga and Anita Govindjee. *Iterative repair planning for spacecraft operations using* *the ASPEN system.* In Artificial Intelligence, Robotics and Automation in Space, volume 440, page 99, 1999. (Cited in page 19.)

[Ramırez 2009] Miquel Ramırez and Hector Geffner. *Plan recognition as planning.* In Proceedings of the 21st international joint conference on Artifical intelligence. Morgan Kaufmann Publishers Inc, pages 1778–1783, 2009. (Cited in pages 19 and 154.)

[Renaudo 2014] Erwan Renaudo, Benoît Girard, Raja Chatila and Mehdi Khamassi. *Design of a control architecture for habit learning in robots.* In Conference on Biomimetic and Biohybrid Systems, pages 249–260. Springer, 2014. (Cited in pages 118, 120, and 180.)

[Renaudo 2015] Erwan Renaudo, Sandra Devin, Benoıt Girard, Raja Chatila, Rachid Alami, Mehdi Khamassi and Aurélie Clodic. *Learning to interact with humans using goal-directed and habitual behaviors.* In RoMan 2015, Workshop on Learning for Human-Robot Collaboration, 2015. (Cited in page 118.)

[Rich 2010] Charles Rich, Brett Ponsler, Aaron Holroyd and Candace L Sidner. *Recognizing engagement in human-robot interaction.* In Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on, pages 375–382. IEEE, 2010. (Cited in page 19.)

[Richardson 2007] Michael J Richardson, Kerry L Marsh, Robert W Isenhower, Justin RL Goodman and Richard C Schmidt. *Rocking together: Dynamics of intentional and unintentional interpersonal coordination.* Human movement science, vol. 26, no. 6, pages 867–891, 2007. (Cited in pages 17 and 153.)

[Rizzolatti 2004] Giacomo Rizzolatti and Laila Craighero. *The mirror-neuron system.* Annu. Rev. Neurosci., vol. 27, pages 169–192, 2004. (Cited in pages 16 and 153.)

[Roberts 2016] Mark Roberts, Ron Alford, Vikas Shivashankar, Michael Leece, Shubham Gupta and David W Aha. *Goal reasoning, planning, and acting with actorsim, the actor simulator.* In ICAPS Workshop on Planning and Robotics (PlanRob), 2016. (Cited in pages 19, 63, and 154.)

[Ros 2010a] Raquel Ros, Séverin Lemaignan, E Akin Sisbot, Rachid Alami, Jasmin Steinwender, Katharina Hamann and Felix Warneken. *Which one? grounding the referent based on efficient human-robot interaction.* In RO-MAN, 2010 IEEE, pages 570–575. IEEE, 2010. (Cited in pages 20 and 154.)

[Ros 2010b] Raquel Ros, Emrah Akin Sisbot, Rachid Alami, Jasmin Steinwender, Katharina Hamann and Felix Warneken. *Solving ambiguities with perspective taking.* In Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction, pages 181–182. IEEE Press, 2010. (Cited in page 45.)

[Roy 2000] Nicholas Roy, Joelle Pineau and Sebastian Thrun. *Spoken dialogue management using probabilistic reasoning.* In Proceedings of the 38th Annual

Meeting on Association for Computational Linguistics, pages 93–100. Association for Computational Linguistics, 2000. (Cited in pages 21 and 154.)

[Salam 2015] Hanan Salam and Mohamed Chetouani. *A multi-level context-based modeling of engagement in human-robot interaction.* In Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on, volume 3, pages 1–6. IEEE, 2015. (Cited in pages 19 and 154.)

[Salem 2011] Maha Salem, Katharina Rohlfing, Stefan Kopp and Frank Joublin. *A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction.* In RO-MAN, 2011 IEEE, pages 247–252. IEEE, 2011. (Cited in page 101.)

[Sanghvi 2011] Jyotirmay Sanghvi, Ginevra Castellano, Iolanda Leite, André Pereira, Peter W McOwan and Ana Paiva. *Automatic analysis of affective postures and body motion to detect engagement with a game companion.* In Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on, pages 305–311. IEEE, 2011. (Cited in pages 19 and 154.)

[Scassellati 2002] Brian Scassellati. *Theory of mind for a humanoid robot.* Autonomous Robots, vol. 12, no. 1, pages 13–24, 2002. (Cited in pages 45 and 160.)

[Sebanz 2006] Natalie Sebanz, Harold Bekkering and Günther Knoblich. *Joint action: bodies and minds moving together.* Trends in cognitive sciences, vol. 10, no. 2, pages 70–76, 2006. (Cited in pages 11, 16, and 152.)

[Sebanz 2009] Natalie Sebanz and Guenther Knoblich. *Prediction in joint action: What, when, and where.* Topics in Cognitive Science, vol. 1, no. 2, pages 353–367, 2009. (Cited in page 16.)

[Shah 2011] Julie Shah, James Wiken, Brian Williams and Cynthia Breazeal. *Improved human-robot team performance using chaski, a human-inspired plan execution system.* In Proceedings of the 6th international conference on Human-robot interaction, pages 29–36. ACM, 2011. (Cited in pages 21, 63, and 154.)

[Singla 2011] Parag Singla and Raymond J Mooney. *Abductive Markov Logic for Plan Recognition.* In AAAI, pages 1069–1075, 2011. (Cited in pages 19 and 154.)

[Sisbot 2012] Emrah Akin Sisbot and Rachid Alami. *A human-aware manipulation planner.* IEEE Transactions on Robotics, vol. 28, no. 5, pages 1045–1057, 2012. (Cited in pages 21 and 155.)

[Skantze 2014] Gabriel Skantze, Anna Hjalmarsson and Catharine Oertel. *Turn-taking, feedback and joint attention in situated human–robot interaction.* Speech Communication, vol. 65, pages 50–66, 2014. (Cited in pages 101, 102, and 174.)

[Smart 2002] William D Smart and L Pack Kaelbling. *Effective reinforcement learning for mobile robots.* In Robotics and Automation, 2002. Proceedings.

ICRA'02. IEEE International Conference on, volume 4, pages 3404–3410. IEEE, 2002. (Cited in page 119.)

[Sorce 2015] Marwin Sorce, Grégoire Pointeau, Maxime Petit, Anne-Laure Mealier, Guillaume Gibert and Peter Ford Dominey. *Proof of concept for a user-centered system for sharing cooperative plan knowledge over extended periods and crew changes in space-flight operations.* In RO-MAN, IEEE, 2015. (Cited in pages 20, 63, and 154.)

[Staudte 2011] Maria Staudte and Matthew W Crocker. *Investigating joint attention mechanisms through spoken human–robot interaction.* Cognition, vol. 120, no. 2, pages 268–291, 2011. (Cited in pages 101 and 102.)

[Stulp 2013] Freek Stulp and Olivier Sigaud. *Robot skill learning: From reinforcement learning to evolution strategies.* Paladyn, Journal of Behavioral Robotics, vol. 4, no. 1, pages 49–61, 2013. (Cited in pages 119 and 179.)

[Sutton 1998] Richard S Sutton and Andrew G Barto. Introduction to reinforcement learning, volume 135. MIT Press Cambridge, 1998. (Cited in page 118.)

[Talamadupula 2014] Kartik Talamadupula, Gordon Briggs, Tathagata Chakraborti, Matthias Scheutz and Subbarao Kambhampati. *Coordination in human-robot teams using mental modeling and plan recognition.* In IROS, IEEE/RSJ, 2014. (Cited in pages 19 and 62.)

[Tambe 1997] Milind Tambe. *Agent Architectures for Flexible, Practical Teamwork.* In 14th National Conf. on AI, 1997. (Cited in page 63.)

[Tenorth 2013] Moritz Tenorth and Michael Beetz. *KnowRob: A knowledge processing infrastructure for cognition-enabled robots.* The International Journal of Robotics Research, vol. 32, no. 5, pages 566–590, 2013. (Cited in page 26.)

[Thomaz 2006] Andrea Lockerd Thomaz, Cynthia Breazeal*et al. Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance.* In Aaai, volume 6, pages 1000–1005, 2006. (Cited in page 119.)

[Thorndike 1911] Edward Lee Thorndike. Animal intelligence: Experimental studies. Macmillan, 1911. (Cited in page 118.)

[Tomasello 2005] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne and Henrike Moll. *Understanding and sharing intentions: The origins of cultural cognition.* Behavioral and brain sciences, vol. 28, no. 05, pages 675–691, 2005. (Cited in pages 12, 13, and 152.)

[Trafton 2013] Greg Trafton, Laura Hiatt, Anthony Harrison, Frank Tamborello, Sangeet Khemlani and Alan Schultz. *Act-r/e: An embodied cognitive architecture for human-robot interaction.* Journal of Human-Robot Interaction, vol. 2, no. 1, pages 30–55, 2013. (Cited in page 26.)

[Tversky 1999] Barbara Tversky, Paul Lee and Scott Mainwaring. *Why do speakers mix perspectives?* Spatial cognition and computation, vol. 1, no. 4, pages 399–412, 1999. (Cited in pages 44 and 159.)

[Venkatesh 2003] Viswanath Venkatesh, Michael G Morris, Gordon B Davis and Fred D Davis. *User acceptance of information technology: Toward a unified view*. MIS quarterly, pages 425–478, 2003. (Cited in page 86.)

[Verbrugge 2008] Rineke Verbrugge and Lisette Mol. *Learning to apply theory of mind*. Journal of Logic, Language and Information, vol. 17, no. 4, pages 489–511, 2008. (Cited in page 44.)

[Vesper 2010] Cordula Vesper, Stephen Butterfill, Günther Knoblich and Natalie Sebanz. *A minimal architecture for joint action*. Neural Networks, vol. 23, no. 8, pages 998–1003, 2010. (Cited in pages 18, 21, 100, 102, 153, and 174.)

[Viejo 2015] Guillaume Viejo, Mehdi Khamassi, Andrea Brovelli and Benoît Girard. *Modeling choice and reaction time during arbitrary visuomotor learning through the coordination of adaptive working memory and reinforcement learning*. Frontiers in behavioral neuroscience, vol. 9, 2015. (Cited in pages 119 and 179.)

[Von Cranach 1973] M Von Cranach and J Heinrich Ellgring. *The perception of looking behaviour*. Social Commun. and Movement, 1973. (Cited in page 100.)

[Waldhart 2016] Jules Waldhart, Mamoun Gharbi and Rachid Alami. *A Novel Software Combining Task and Motion Planning for Human-Robot Interaction*. In 2016 AAAI Fall Symposium Series, 2016. (Cited in page 30.)

[Walters 2007] Michael L Walters, Kerstin Dautenhahn, Sarah N Woods and Kheng Lee Koay. *Robotic etiquette: results from user studies involving a fetch and carry task*. In Proceedings of the ACM/IEEE international conference on Human-robot interaction, pages 317–324. ACM, 2007. (Cited in pages 21 and 155.)

[Warnier 2012] Mathieu Warnier, Julien Guitton, Séverin Lemaignan and Rachid Alami. *When the robot puts itself in your shoes. managing and exploiting human and robot beliefs*. In RO-MAN, 2012 IEEE, pages 948–954. IEEE, 2012. (Cited in pages 19, 47, 62, and 160.)

[Watkins 1989] Christopher John Cornish Hellaby Watkins. *Learning from delayed rewards*. PhD thesis, University of Cambridge England, 1989. (Cited in page 121.)

[Weistroffer 2014] Vincent Weistroffer. *Étude des conditions d'acceptabilité de la collaboration homme-robot en utilisant la réalité virtuelle*. PhD thesis, Paris, ENMP, 2014. (Cited in pages 87 and 171.)

[Woods 2006] Sarah Woods, Michael Walters, Kheng Lee Koay and Kerstin Dautenhahn. *Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach*. In Advanced Motion Control, 2006. 9th IEEE International Workshop on, pages 750–755. IEEE, 2006. (Cited in page 103.)

[Zaga 2017] Cristina Zaga, Roelof AJ de Vries, Jamy Li, Khiet P Truong and Vanessa Evers. *A Simple Nod of the Head: The Effect of Minimal Robot*

Movements on Children's Perception of a Low-Anthropomorphic Robot. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pages 336–341. ACM, 2017. (Cited in page 101.)