



HAL
open science

Human Aware Task Planning Using Verbal Communication Feasibility and Costs

Guilhem Buisan, Guillaume Sarthou, Rachid Alami

► **To cite this version:**

Guilhem Buisan, Guillaume Sarthou, Rachid Alami. Human Aware Task Planning Using Verbal Communication Feasibility and Costs. 12th International Conference on Social Robotics (ICSR 2020), Nov 2020, Golden, United States. pp.554-565. hal-02927562

HAL Id: hal-02927562

<https://hal.science/hal-02927562v1>

Submitted on 7 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Human aware task planning using verbal communication feasibility and costs

Guilhem Buisan¹[0000-0001-6277-9820], Guillaume Sarthou¹[0000-0002-4438-2763],
and Rachid Alami^{1,2}[0000-0002-9558-8163]

¹LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France

²Artificial and Natural Intelligence Toulouse Institute (ANITI)

{guilhem.buisan, guillaume.sarthou, rachid.alami}@laas.fr

Abstract. This paper addresses the problem of taking into account the feasibility and cost of verbal communication actions at the task planning level in the context of human-robot interaction. By determining the human and robot context in which the referring expressions to entities can be effectively achieved and at which cost, our planner can 1) prevent potential deadlocked situations where it is not possible to produce a feasible communication act 2) find plans reducing the overall communication complexity 3) evaluate different communications strategies and select the most suitable one. Our approach is based on the extension of a multi agent hierarchical task planner capable of maintaining, throughout the planning process, an estimate of the human partner knowledge about the environment using a semantic representation. A Referring Expression Generation (REG) can then be executed for each communication action allowing the task planner to be informed about its feasibility and cost during the plan elaboration process. Three scenarios are presented to validate the method, where a robot drives the human actions to perform a cube arrangement task.

Keywords: Human-Robot Collaboration · Hierarchical Task Planner · Referring Expression Generation · Verbal Communication Planning

1 Introduction

It is well established that clear and fluent communication is a key aspect of the success of collaborative tasks. In the Human-Robot Interaction (HRI) research field it has lead to two dual problems [17]. In one hand we have the language understanding in which the robot interprets and grounds human’s utterances and reacts to them [1]. In another hand we have the language generation problem where the robot produces language whether to ask for help [18], align knowledge [3] or clarify its decision [11]. In this paper, we only use the language generation and consider it as a communication action that is part of a Human-Robot collaborative task plan. Taking it as an action means that the robot can plan when and what to communicate. More importantly, by taking into account the *what* of the communication, its content, it is possible to estimate the feasibility and cost of a

communication action. In addition, by maintaining a representation of the environment for the future states of the task as it is done in symbolic task planning, the robot can estimate the feasibility and the cost of the verbal communication actions all along the task. Considering these two pieces of information in the planning process makes it possible to compare verbal communication with other means of communication, to find a plan minimizing the overall communication complexity but also to prevent some plan failures. This process can be compared to [7] where the geometric feasibility, the indirect effects and cost of an action is considered already during the symbolic task planning and not at a further geometric planning level.

Consider the situation illustrated in Fig. 1. In this task the human and the robot have to put each cube in a specific colored area. The robot does not have the ability to act on the cubes and the human does not know the goal configuration. The robot must therefore communicate the successive actions that the human will have to perform. The two cubes are visually the same to the human, but the robot can identify them. The initial configuration is given in Fig. 1a with the cube C1 in the red area and the cube C2 in the black area. The goal configuration (Fig. 1d) requires the cube C1 to be placed in the black area and the cube C2 in the white area. In the case where the communication action is refined only at execution, a solution plan could be to tell the human to move the cube C1 in the black area then the cube C2 in the white one. The execution of this plan would result in: *"Take the cube in the red area and put it in the black area"*. In this new situation where both cubes are now in the black area (Fig. 1b), the robot has no way to designate the cube C2 without ambiguity. Hence, the task is blocked. Taking into account the communication feasibility and cost estimation during the planning process would allow to find the solution where the robot tells the human to move the cube C2 first (Fig. 1c) and then the cube C1 (Fig. 1d). Considering now that the robot can point to the cubes, the deadlock of the first solution can be avoided with a pointing action and nevertheless, thanks to the communication cost estimation, the least expensive solution can be selected.

In §2, we briefly discuss related work and how our contribution addresses new issues. Our approach and its components are then described in §3. Three case studies are finally presented in §4 to show how this approach can be used to prevent deadlocked situations at execution, how it can reduce the global communication complexity during a Human-Robot collaborative task and how it can be used to balance between different communication means.

2 Related Work

A significant amount of research has been dedicated to Human-Robot verbal communication, especially to answer the questions of *what* and *when* to communicate [9]. A lot of early works address these questions at execution time, with a fixed plan in which the robot inserts verbal communication when needed.

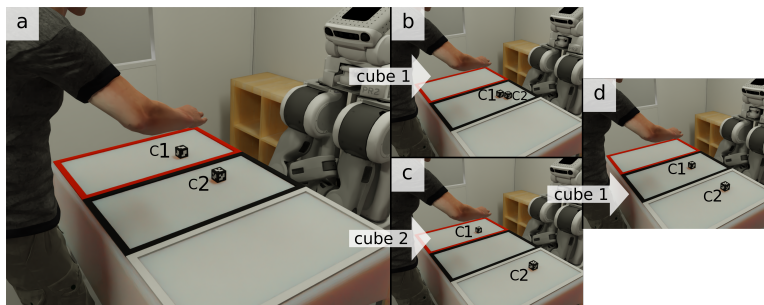


Fig. 1. A Human-Robot collaborative task where the robot has to explain to the human to put the cube C1 in the black area and the cube C2 in the white area. The cube identifiers ("C1", "C2") are only known to the robot. If the action of referring to an object without ambiguity is not taken into account during task planning, a deadlocked situation could appear if the robot first asks to move the cube C1.

The communication can be used to share the plan, to ask for or give specific information or to fix errors ([3], [14], [15], [16], [18]).

Devin and Alami [3] use a theory-of-mind enabled framework to decide, at execution time, if a communication action is needed and its contents. In their work, the robot is provided with a shared plan for both itself and the human partner. During the execution, for each human action, the robot estimates the partner's mental model. Divergence between the robot knowledge and the estimated human knowledge is monitored. When a divergence which can endanger the plan is detected, a verbal communication of the needed facts is computed and achieved.

However, in some cases, deciding on communication at execution time is not enough and more recent work deals with explicit consideration of communication actions while planning ([10], [11], [12]). Roncone et al. [11] explicitly represent three types of verbal communication action, *command* (instruction to the human) which can be declined, *ask* (a question of whether her current action is over) and *inform* (the communication of a robot intent). These actions are integrated with others classical actions into a Partially Observable Markov Decision Process (POMDP) which, once solved, returns a policy integrating communication actions. However, the costs associated with these communication are not taken into account for task allocation. Moreover, the content of the communication (the *what*) is not specified at planning level, which can cause unplanned higher communication costs or even non achievable communication in some scenarios.

A similar approach is proposed by Unhelkar et al. [19] with multiple verbal communication types considered (*command*, *ask*, *inform* and *answer*). Moreover a communication cost is explicitly considered but it is a cost on the *when* to communicate and not on the *what*, as it is a function penalizing temporally close communication actions. The communication actions are task specific and include parameters replaced at execution time (the landmarks and actions for the *inform* communication in their example; *i.e.* the *what* of the communication). In their

examples, every landmark is assumed to be easily referred to the human, but this is not always the case. By using REG at planning time, our approach addresses two of the five challenges identified by Unhelkar et al.: "estimating benefit of communication" and "quantifying cost of communication" [20].

To entrench the aim of this paper, we will use the example depicted by Tellex et al. [18], and describe how our approach could differ from theirs. In this example, a robot and a human are engaged in assembling tables, the robot is following a precomputed plan. When a failure occurs the robot asks the human for help by referring to problematic objects. By doing so, the robot plan is repaired with the human help thanks to an effective object referring communication action. However, in this work, they consider a non-reachability of a table leg as a failure even if this table leg was known by the robot to be unreachable from the beginning. A planning step beforehand could assign the assembly of this leg to the human, and plan the needed communication of the robot to the human. Moreover, with the approach presented in this paper, if multiple table legs would match the verbal description, the planner would be able to insert an action or modify the order of assembly to make the needed referring easier to understand for the human.

3 Method

In this section, we first provide an overview of our approach and briefly describe the used Hierarchical Agent-based Task Planner [8]. We then introduce the reader to REG, based on [2], and the knowledge base it needs. We end this section with a presentation of the integration of the two components and how it allows the planner to be informed of the feasibility and the cost of the communication actions.

3.1 Approach

The communication actions that we consider in this paper are commands issued by the robot based on Referring Expressions (REs). Typical commands are "*Take X*" and "*Put it in Y*". We thus have a static part and the rest depends on the situation when the communication is performed and must be solved by a REG. It is this variable part that could make a communication costly or infeasible. As it has been highlighted in [2], REG must be performed on the human's partner Knowledge Base (KB) to only use facts and concepts that the robot estimates to be known by the human. Because of that, we target a planner that is already suitable for HRI to integrate the estimation of communications. This means that we need a planner able to distinguish between the different agents involved in the task and to maintain a representation of the environment for each of them.

Because the task the planner has to solve does not necessarily imply all the elements present in the current environment, the planner does not need a full representation of the environment. In a same way, it does not necessarily need to have all the characteristics of the entities such that their colors or their types. On the example of Fig. 1, the two cubes can only be represented as movable

objects in the planner and not as cubes to make the planning domain more generic. However, the REG needs all the semantic information of each entity of the environment to generate accurate RE. Furthermore, if another cube which is not part of the task, thus not part of the planner internal representation, is present on the table, it will also impact the REG and thus the complexity and feasibility of the communication action. Hence, the REG can not be performed on the planner internal representation. To solve this issue, we endow the planner with the ability to maintain a semantic KB that is used by the REG. Since maintaining this external representation can be an heavy process, it is updated only when a communication action has to be evaluated.

The general workflow executed for each communication action encountered during the planning process consists of: 1) updating the external semantic KB of the human partner with the expected world state 2) identifying the objects to which to refer to in the communication 3) execute the REG for each of these objects 4) calculate the feasibility and the cost of the communication action according to the feasibility and the cost of each individual RE involved in the planned communication. Note that the examples used in this paper only involve one RE but the same method can be used for communications of type "give me X and Y". In this case, the external semantic KB is only updated once and both REG are executed on this KB.

3.2 Hierarchical Task Planner

In order to implement our approach, we need a task planner able to maintain an estimated knowledge base of each agent at each planning step. We chose the Hierarchical Agent-based Task Planner (HATP) [8]. HATP extends the classical Hierarchical Task Network (HTN) planning by being able to produce *shared plans* to reach a joint goal. A HATP planning domain describes how to decompose tasks into subtasks down to atomic symbolic actions. Both the robot and human feasible tasks and actions are described in the domain. A context-dependent cost function is associated to each action.

During the task decomposition, HATP will explore several applicable subtasks until the global task is totally refined into feasible actions, and will return the minimal cost plan. HATP also supports *social rules*, allowing to balance the effort of involved agents depending on human preferences and to penalize plans presenting certain undesirable sequences of actions. We will not use these social rules in what follows, but our approach stays totally compatible with them.

Moreover, during the exploration of the task tree, HATP will assign actions to available agents, robot or human (when an action can be done by both). By doing so, HATP is able to elaborate one action *stream* per agent, together with causality and synchronization links. Besides, HATP domain syntax supports Multiple Values State Variables (MVSV) [6] which is used to represent and reason about each agent mental state. The value of each variable depends on the agent it is requested for. This allows to represent action preconditions depending on the knowledge of the agent performing the action and also represent their effect on each agent mental state which can dependent on the agent perspective.

Finally, the last argument which motivated our choice was the previous integration of HATP with a Geometrical Task Planning (GTP) [5]. This work aimed at refining geometric and motion planning requests during the task planning process. The geometric planner would then compute, in context, the feasibility, the cost and the side effects of the action. In a similar way, we propose here to integrate and run REG, in context, to determine communication action feasibility and pertinence with respect to other courses of actions.

3.3 Referring Expression Generation

The REG aims to unambiguously designate an entity/object a_t in an environment. We use the state of the art algorithm presented in [2]. This REG runs on a semantic knowledge base K represented as an ontology. In its simplified version, we defined $K = \langle \mathbb{A}, \mathbb{T}, \mathbb{R} \rangle$ where \mathbb{A} , \mathbb{T} and \mathbb{R} are respectively an ABox, TBox and RBox [4]. \mathbb{T} is the TBox of the ontology defining a set of classes and a finite collection of class inclusion axioms. RBox is a set of properties and a finite collection of property inclusion axioms, transitivity axioms, reflexive axioms, inverses axioms and chain axioms. \mathbb{A} , the ABox of the ontology, is a set of entities A and a finite collection of axioms of the form (a_i, p, a_j) and (a_i, t) , where a_i and a_j are entities, t is a class and p is a property. All these axioms can be represented as relations of the form of triplet $r = (s, p, o)$ with respectively s , p and o the subject, property and object of the relation r . The axiom (a_i, t) corresponds to the inheritance relation that is written $r = (a_i, isA, t)$. This means that entities of ABox are instances of TBox classes. To manage the ontologies, we use Ontogenius [13] which is the ontology manager on which the REG we use was developed.

The REG problem is defined as a tuple $\mathcal{P} = \langle a_t, K, Ctx \rangle$. In this problem, a_t is the target entity. The one to be referenced through its relations to other entities existing in the ontology K . Ctx is the context of the problem. It is a set of relations considered as already known by the communicating agents about a_t . It is therefore a subset of \mathbb{A} . With the table-top interaction of this paper, the context is $Ctx = \{(a_t, isOn, table_1)\}$ because both interacting agents are aware that they are currently speaking about the objects on the table. This means that the REG has not to consider the objects not being on the table as possibly bringing ambiguity. The REG context could be enriched by taking for example the visibility of the human partner and should be determined dynamically during the task planning process. However, for the demonstration of our current purpose, we consider it fixed.

A solution to the REG problem is a set of relations which could be verbalized afterwards and a cost associated with this solution. The cost of the solution is the sum of the cost of each property used in the set of relations. These costs can differ from one human to another and represent the effort needed by the human to interpret the relation. The fact that the property "hasColor" would be easier to interpret than the property "isAtLeftOf" would be represented by a higher cost for the "isAtLeftOf" property.

3.4 Integration of REG within Action Planning

The representation of the communication action: For clarity purposes, in this paper, we only place ourselves in scenarios where only the robot knows the goal of a joint task and issues command to its human partner one at a time when the human has to do an action. Thus, while planning, if a task is allocated to the human, as she has no way of guessing it, a preceding communication is required. In the HATP domain, this translates as a method being decomposed into a sequence of a communication action and an action made by the human when the task is attributed to the human. The communication action feasibility is determined by both symbolic preconditions (*e.g.* the human and the robot are in the same room) and REG result (whether a solution is found or not). If the communication action is feasible, the cost of the communication action is then computed as the sum of a fixed cost depending on the type of communication and the REG solution cost depending on the human receiver and the entities to refer to in the communication.

We have chosen here for illustration purposes a simple planning problem where a communication needing a REG is involved in each plan step, but the method is general and compatible with problems which need to estimate and ensure the pertinent context and plan step (the *when*) of a communication action during plan elaboration (*e.g.* [3], [19]).

Maintaining the right knowledge base, at the right time: On one hand, we have large, complete semantic knowledge bases on which a REG algorithm is able to run and to return the feasibility, the cost and the content of a verbal entity referring communication for a specified agent (top part of Fig. 2). On the other hand, we have reduced knowledge bases dedicated to task planning (bottom part of Fig. 2). In order to know the feasibility and cost of a verbal communication action during the planning, we have to reconcile both sides. Indeed, the estimated ontology of the communication receiver must be updated to reflect her planned estimated beliefs at the time of the communication. All the knowledge representation used here are from the robot point-of view and managed internally by the robot decisional and knowledge management processes.

First, the attributes of all the entities present in the planning knowledge base are initialized for each agent (left part of Fig. 2). To do so, every entity types declared in the planning domain are retrieved from the ontologies by their name, and entities inheriting from these types in the ontologies are created in the planning knowledge base. Then, each attribute (both static and dynamic) of every entities declared in the domain has its value updated. If the attribute is a set, multiple relations with the same name originating from the same entity and pointing to different ones can be found in the ontologies. If so, all the pointed entities are added to the set. Finally, a planning ontology is created by copy of the present one for every agent other than the robot present in the planning domain. These copies are made to avoid modifying the original ontologies during the planning process as other components may rely on them.

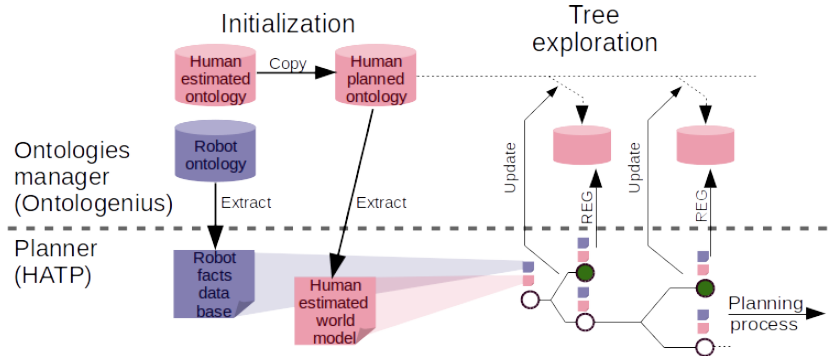


Fig. 2. A representation of the exploration of potential mental states and ontologies conducted by the planner. The ontology representing estimated human knowledge is first copied in order to plan it without altering the original one. The human and robot planning information is extracted from the ontologies. During the tree exploration, for each verbal communication action, the planned human ontology is updated with the current explored state and the REG is executed on it.

When a communication action is encountered during the task tree exploration (right part of Fig. 2), the ontology of the communication receiver needs to be updated to be able to run the REG on it. The planning ontology copy of the receiver human is retrieved by her identifier. Then, for each of the entities of her planned beliefs at the time of the communication, an update is made. The update is only made on dynamic attributes as static ones do not change during the planning process. All the relations having the same name as the attribute of the entity in the planning domain are deleted from the ontology, and replaced with new planned values. If the attribute is a set, a new relation with the same name is created for every value in the set.

A REG request is then issued on the updated ontology with the goal individual being the entity to refer. The REG returns a solution with a cost or a failure which is taken into account by the planner as classical cost or a non fulfillment of the action preconditions respectively. Alternatively, a communication action may need to refer to multiple entities. In that case, multiple REG requests are issued on the same updated ontology and their costs are summed.

4 Case Studies

In this section, we present three case studies. The two first ones are run in simulation on a minimalist setup and show respectively that the estimation of the communication content during the planning can prevent from execution dead-end and can reduce the global communication complexity during the task. The third case study is run on a PR2 robot with a perception of its environment and presents a more complex task with twelve objects to organize. With this last

case, we show that our method makes it possible to compare different means of communication and to choose the most appropriate.

All the three cases studies are based on a cube arrangement task. The human can distinguish the cubes by their color and the digit written on them (one or two) if there is one. As shown in Fig. 1, the table surface is composed of three storage areas of different colors and cubes can be placed only in one of them. This position information can also be used by the human to distinguish the cubes. By this way, the robot can refer to a cube with a REG of the type: *"the black cube with the number 2 which is in the black area"*. In all the cases, only the robot knows the goal position the cubes but can not manipulate them. It thus has to guide the human in the arrangement task. The robot can only point the cubes in the third case study. In the first two, he can only use verbal communication.

4.1 Preventing execution dead-ends

In this case study, we consider the initial state presented in Fig. 1. The cube C1 is in the red area and the cube C2 in the black one. The goal state is to have the cube C1 in the black area and the cube C2 in the white one (Fig. 1d). Taking into account the cost and the feasibility of the communication, we found with our method the plan 1. Cube C2 is moved first because otherwise the two cubes would be in the black area at the same time. Such a situation would cause a dead-end during the execution of the plan or require another communication mean.

```

HR - TellHumanToTake(C2) // (C2, isA, Cube), (C2, isIn, AB),
                          // (AB, isA, Area), (AB, hasColor, black)
H  - Take(C2)
HR - TellHumanToPlace(C2, AW) // (AW, isA, Area), (AW, hasColor, white)
H  - Place(C2, AW)
HR - TellHumanToTake(C1) // (C1, isA, Cube), (C1, isIn, AR),
                          // (AR, isA, Area), (AR, hasColor, red)
H  - Take(C1)
HR - TellHumanToPlace(C1, AB) // (AB, isA, Area), (AB, hasColor, black)
H  - Place(C1, AB)

```

Plan 1. The obtained plan for the first case study where cube C1 must be moved from the red to the black area and cube C2 moved from the black to the white area. The lines beginning with H represent the actions of the human and the lines beginning with HR represent actions involving the human and the robot (communication actions). In green are the REG results for each communication action.

We consider once again the initial state presented in Fig. 1. This time the goal is to invert the positions of the two cubes. In this situation, if the communication cost and feasibility is not taken into account during planning, both actions directly leading to the goal state (*i.e.* cube C1 moved to black area or cube C2 to red area) will lead to a dead-end at plan execution. The solution found with our method is to add a supplementary action. It consists in putting the cube C1 away (in the white area). This additional action avoids a dead-end by making communication about cube C2 feasible.

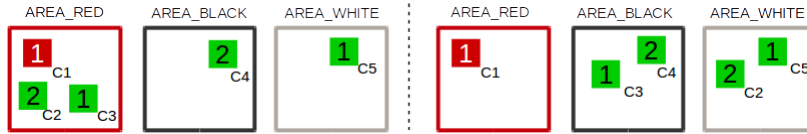


Fig. 3. The initial state (left) and the goal state (right) of a task where the robot has to explain to the human partner how to move the cubes to complete the task.

4.2 Reduction of the overall communication complexity

In this second case study, we show how estimation of communication by verbal designation can be used to reduce the complexity of global communication. This time we consider the initial state and the target state represented in Fig. 3. Only cubes C2 and C3 should be moved. Our method finds the solution consisting in moving cube C2 first, then cube C3. With this order, cube C2 is referred by three relations: its type (*i.e.* cube), the number on it and the colored area in which it is located. After that, the cube C3 can also be referred only by three relationships being its type, its color and the colored area in which it is located. Considering the reverse order, this would have generated a more complex RE first for cube C3 with four relationships: its type, its color, the number on it and the colored area in which it is located. The solution chosen by our method communicates a sum of six relations rather than seven with the reverse order.

4.3 Compare with other communication means

In this last case study, we show how the estimation of verbal designation communication cost can be used to compare it with other communication means, here pointing. Now, we consider twelve cubes. The initial state and the goal state are represented on Fig. 4. Such a number of similar objects leads to long explanations to refer to certain cubes. Therefore, we aim the task planner to choose another means of communication to refer to these cubes (*e.g.* a pointing action). The pointing action has a constant cost which is higher than a simple verbal communication but lower than a complex one (with three or more relations to verbalize). To exemplify the comparison with other communication means, the arrangement order is predefined in this setup.

The execution of the computed plan can be found in the video available at https://youtu.be/3YnGh_t-UpY. The cubes C5 and C7 are chosen to be pointed instead of verbalized. Indeed, in the world states where these cubes need to be moved, verbal referring is considered to be too costly, thus a pointing motion is preferred. For example, the cube C5 in the initial situation needs a long and complex explanation that is: *"take the black cube with the number two which is in the black area"*. Even in the case the pointing action takes more execution time, it could require less cognitive load for the human partner and so make the human action faster.

Here, we see another benefit of our approach, it allows the planner to balance between the use of verbal communication actions, which can become complex in

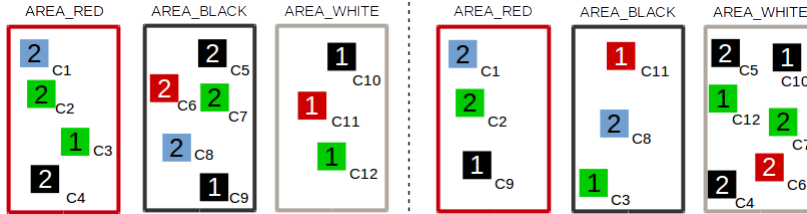


Fig. 4. The initial state (left) and the goal state (right) of a task where the robot has to explain to the human partner how to move the cubes to complete the task.

some states (hard to predict without a task planner), and other communication modalities. Here, verbal communication is balanced with other communication means, but it can also be balanced with other actions or assignments requiring less or no communication.

5 Conclusion and future work

This paper presents two main contributions. The first contribution deals with the integration of a fully informed refinement of communication actions in the shared plan elaboration process. Indeed, dealing with the refinement of communication actions in a subsequent planning step or postponing it to the execution phase might entail inefficiencies or even plan failures. The second contribution concerns the interplay between the exploration and evaluation of future potential states by the task planner and the management of ontological and semantic knowledge both for the robot and its estimated knowledge of the human partner.

The result is an extended version of HATP, called HATP/REG. Its abilities have been illustrated in three case studies showing its ability 1) to prevent potential deadlocked situations where it is not possible to produce a feasible communication act (and specifically here a non ambiguous reference) 2) to find plans reducing the overall communication cost 3) to balance between different communications means and select the most suitable one.

In the current version, the refinement focuses on the referencing problem of objects and action parameters. We intend to extend it to other communication acts such as plan and/or sub-plan explanation. Finally, we also intend to integrate other communication means in the framework such as more complex pointing actions and more elaborate mental states (*e.g.* integrating the shared human-robot experience). We will also conduct an evaluation of the system through user studies to assess its pertinence.

Acknowledgements

We thank Amandine Mayima for her contribution to the implementation of the integrated scenarios. This work has been funded by the French Agence Nationale de la Recherche JointAction4HRI project ANR-16-CE33-0017.

References

1. Brawer, J., Mangin, O., Roncone, A., Widder, S., Scassellati, B.: Situated human-robot collaboration: predicting intent from grounded natural language. In: IEEE/RSJ IROS (2018)
2. Buisan, G., Sarthou, G., Bit-Monnot, A., Clodic, A., Alami, R.: Efficient, situated and ontology based referring expression generation for human-robot collaboration. In: RO-MAN. IEEE (2020)
3. Devin, S., Alami, R.: An Implemented Theory of Mind to Improve Human-Robot Shared Plans Execution. In: ACM/IEEE HRI (2016)
4. Fokoue, A., Kershenbaum, A., Ma, L., Schonberg, E., Srinivas, K.: The Summary Abox: Cutting Ontologies Down to Size. In: The Semantic Web - ISWC, vol. 4273. Springer Berlin Heidelberg (2006)
5. Gharbi, M., Lallement, R., Alami, R.: Combining symbolic and geometric planning to synthesize human-aware plans: toward more efficient combined search. In: RSJ IROS. IEEE (2015)
6. Guitton, J., Warnier, M., Alami, R.: Belief management for hri planning. In: European Conf. on Artificial Intelligence (2012)
7. Lallement, R.: Symbolic and Geometric Planning for teams of Robots and Humans. Theses, INSA de Toulouse (2016)
8. Lallement, R., De Silva, L., Alami, R.: HATP: An HTN Planner for Robotics. In: ICAPS Workshop on Planning and Robotics (2014)
9. Mavridis, N.: A review of verbal and non-verbal human-robot interactive communication. *Robotics and Autonomous Systems* **63** (2015)
10. Nikolaidis, S., Kwon, M., Forlizzi, J., Srinivasa, S.: Planning with verbal communication for human-robot collaboration. *ACM Transactions on Human-Robot Interaction (THRI)* **7**(3) (2018)
11. Roncone, A., Mangin, O., Scassellati, B.: Transparent role assignment and task allocation in human robot collaboration. In: ICRA. IEEE (2017)
12. Sanelli, V., Cashmore, M., Magazzeni, D., Iocchi, L.: Short-term human-robot interaction through conditional planning and execution. In: ICAPS (2017)
13. Sarthou, G., Clodic, A., Alami, R.: Ontologenius: A long-term semantic memory for robotic agents. In: RO-MAN. IEEE (2019)
14. Schaefer, K.E., Straub, E.R., Chen, J.Y., Putney, J., Evans III, A.W.: Communicating intent to develop shared situation awareness and engender trust in human-agent teams. *Cognitive Systems Research* **46** (2017)
15. Sebastiani, E., Lallement, R., Alami, R., Iocchi, L.: Dealing with On-line Human-Robot Negotiations in Hierarchical Agent-Based Task Planner. In: ICAPS (2017)
16. Shah, J., Wiken, J., Williams, B., Breazeal, C.: Improved human-robot team performance using chaski, a human-inspired plan execution system. In: ACM/IEEE HRI (2011)
17. Tellex, S., Gopalan, N., Kress-Gazit, H., Matuszek, C.: Robots that use language. *Annual Review of Control, Robotics, and Autonomous Systems* **3**(1) (2020)
18. Tellex, S., Knepper, R., Li, A., Rus, D., Roy, N.: Asking for help using inverse semantics. *Robotics: Science and Systems X* (2014)
19. Unhelkar, V.V., Li, S., Shah, J.A.: Decision-making for bidirectional communication in sequential human-robot collaborative tasks. In: ACM/IEEE HRI (2020)
20. Unhelkar, V.V., Yang, X.J., Shah, J.A.: Challenges for communication decision-making in sequential human-robot collaborative tasks. In: Workshop on Mathematical Models, Algorithms, and Human-Robot Interaction at RSS (2017)