



HAL
open science

Towards multi-state visuo-spatial reasoning based proactive human-robot interaction

Amit Kumar Pandey, Muhammad Ali, Matthieu Warnier, Rachid Alami

► To cite this version:

Amit Kumar Pandey, Muhammad Ali, Matthieu Warnier, Rachid Alami. Towards multi-state visuo-spatial reasoning based proactive human-robot interaction. 15th International Conference on Advanced Robotics (ICAR), Dec 2011, Tallinn, Estonia. hal-01977501

HAL Id: hal-01977501

<https://laas.hal.science/hal-01977501>

Submitted on 10 Jan 2019

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.

Towards Multi-State Visuo-Spatial Reasoning based Proactive Human-Robot Interaction

Amit Kumar Pandey, Muhammad Ali, Matthieu Warnier and Rachid Alami

Abstract—Robots are expected to co-operate with humans in day-to-day interaction. One aspect of such co-operation is behaving proactively. In this paper, our robot will exploit the visuo-spatial perspective-taking of the human partner not only from his current state but also from a set of different states he might attain from his current state. Such rich information will help the robot in better predicting ‘where’ the human can perform a particular task and how the robot could support it.

We have tested the system on two different robots for the tasks of *giving* and *making an object accessible* to the robot by the human partner. Our robots equipped with such *multi-state visuo-spatial perspective-taking capabilities* show different proactive behaviors depending upon the task and situation, such as reach out proactively and to a correct place, when human has to give an object to the robot. Primary results of user studies show that such proactive behaviors reduce the human’s ‘confusion’ as well as ‘the robot’ seems to be more ‘aware’ about the task and the human.

I. INTRODUCTION

AS robots are moving towards being co-operative and social, the challenges of incorporating the basic ingredients of such behaviors are becoming more prominent. Behaving proactively is one of such desirable characteristics. Proactive behavior could be at various levels of abstractions as well as in various ways ranging from simple verbal interaction [11, 12] to proactive task selection [1-4, 5, 7, 19].

In this paper our focus will be on the situations where human is required to perform a task for the robot. In this context we identify that for behaving proactively, robot needs to predict (i) ‘what’ the human is intended to do? (ii) ‘When’ the human can or want to do it? (iii) ‘How’ the human might do it? (iv) ‘Where’ the human might or might not perform it. Study in [8] shows human’s abilities of predicting ‘what’, ‘when’ and ‘where’ of others actions and its usefulness in online action coordination, in the context of joint actions.

Predicting ‘what’ has been addressed in various ways. Robot estimates what human wants and selects a task using probability density function, [1, 5], whereas in [7] it performs anticipatory action selection for improving joint task coordination. Predicting ‘when’ and ‘how’ have also been addressed in various ways: temporal Bayesian

networks for proactive action selection to minimize wait time, [4], robot wheel chair taking control when human needs it, [17], activity constraints violation based scheduler to remind human, [12], using switching hidden semi-Markov model to learn house occupant’s daily activities and to alert the caregiver in case of abnormality, [18].

On the other hand the domain of predicting ‘where’ the human can perform the task and to show proactive behavior accordingly, has not been explored enough in robotics. In [16], robot wheel chair, using POMDP, predicts where the human is trying to go and drives him there.

Predicting where human can perform a task is helpful for sharing attention with others [9], to predict spatial characteristics of the others actions [10], for building robot’s theory of mind [13] and consequently could help guide human behavior towards robot, [14, 15].

In this paper we will address this issue of predicting the places ‘where’ human can perform a task. The task might have been predicted by the robot or already known to the robot. For example, if human wants to give some object to the robot, the robot will predict where human can perform that task as well as where robot can take the object, then it will move its hand proactively to support the task as well to guide the human to achieve the joint task of hand over. Such behavior will also make the human ‘aware’ about the ‘awareness’ of the robot for the current task. As shown in fig. 1 (a), the robot requested, “Please give me the toy dog.” and remained in the rest position. This might create confusion for human about how and where to give it to the robot: “Should I move and reach near to the robot or should I stand and put the object in the hand of the robot or should I put it at a place somewhere on the table for the robot to take it?” But if the robot along with the request to give, also moves its hand towards probable place to take, as shown in fig 1(b), it will greatly guide the human about how and

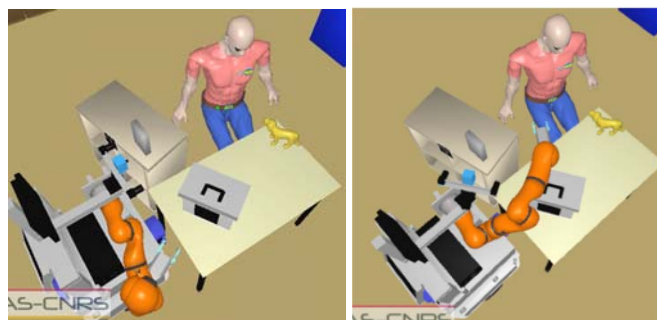


Fig. 1: The Robot asks to the human “give me the toy dog” (a) by maintaining its rest position, (b) by proactively moving its hand to appropriate place to guide human’s action.

Authors are with CNRS; LAAS; 7 avenue du Colonel Roche, F-31077 Toulouse, France. Université de Toulouse; UPS, INSA, INP, ISAE; LAAS; F-31077 Toulouse, France.

Email: {akpandey, mali, mwarnier, rachid.alami}@laas.fr

The work described here has been partially conducted within the EU Project CHRIS (Cooperative Human Robot Interaction Systems) funded by the E.C. Division FP7-IST under Contract 215805.

TABLE I
STATES FOR MULTI-STATE VISUO-SPATIAL PERSPECTIVE TAKING

Reachability States	Visibility States
Current	Current
Sitting Straight	Sitting Straight Head
Sitting Turn Around	Sitting Turn Head
Sitting Lean Forward	Sitting Lean Torso and Turn Head
Sitting Turn and Lean	Sitting Turn Torso and Turn Head
Standing Straight	Sitting Turn-Lean Torso and Turn Head
Standing Turn Around	Standing Straight Head
Standing Lean Forward	Standing Turn Head
Standing Turn and Lean	Standing Lean Torso and Turn Head
	Standing Turn Torso and Turn Head
	Standing Turn-Lean Torso and Turn Head

where to give. In this context it is important for the robot to be aware about human’s capabilities and move its hand to appropriate place, which reduces the human’s effort as well. Moreover, robot moving its arm to a location will induce human goal anticipatory response as demonstrated in [6].

To facilitate such reasoning our robot will use visuo-spatial perspective taking of the human. The key is, it does such reasoning not only from his current state but also from a set of different states the human might attain. Furthermore it performs similar analysis for itself.

In the next section we will first briefly describe the different states for which robot estimates visuo-spatial abilities of the agent. Then we will briefly explain the robot’s understanding of a task from the visuo-spatial perspective. Then the task dependent analysis of the common or mutually exclusive abilities will be explained to predict the probable places for human to perform the task and assign an intuitive level of comfort. Then we explain different proactive behaviors depending on the task and the scenario. Then the conclusion and pointer will follow the experimental results and analysis to future works.

II. METHODOLOGY

A. Multi-state visuo-spatial analysis

As explained earlier, our robots are capable of analyzing various abilities of an agent not only from the current position of the agent but also from a set of different possible positions the agent might attain. Table I shows the list of different states for calculating reachability and visibility of an agent. Such abilities are calculated in terms of cells of the discretized 3D workspace and at objects’ level in workspace, which we termed as *Mightability Maps (MM)*, [20] and *Object Oriented Mightabilities (OOM)* respectively. Such analyses enable the robot to reason about: if the agent will stand up and lean forward, which are the visible and reachable places and objects, and so on.

B. Categorizing efforts

As the robot is having a rich multi-state visuo-spatial

TABLE II
EFFORT CLASSES FOR VISUO-SPATIAL ABILITIES

Effort to Reach	Effort Level	Effort to See
No_Effort_Required	 Minimum	No_Effort_Required
Arm_Effort		Head_Effort
Arm_Torso_Effort		Head_Torso_Effort
Whole_Body_Effort		Whole_Body_Effort
Displacement_Effort		Displacement_Effort
No_Possible_Known_Effort	Maximum	No_Possible_Known_Effort

perception of the agent’s ability, we categorize efforts to attain one state from the current state in terms of the joints involved in such transitions. Table II shows such categorization. It is motivated from the studies of human movement and behavioral psychology, [21], where different types of reach actions of the human have been identified and analyzed, as shown in fig. 2, reach involving simple arm extension (arm-only reach), shoulder extension (arm-and-shoulder reach), leaning forward (arm-and-torso reach) and standing reach. Such categorization could be further enhanced based on the studies of musculoskeletal kinematics and dynamics models such as [23].

C. Robot’s Understanding of a task

In [25] we have presented an approach, which facilitates the robot to understand a task from human-human demonstration. The idea is that as the robot understands task semantics from the point of view of desirable effects in terms of visuo-spatial abilities, as an attempt to make the understanding independent of how to perform the task. The agent for whom the task is being performed, we termed it as ‘*target-agent*’. As such understanding will partially serve our robot to decide the way to behave proactively, in table III we have summarized the robot’s understanding for a set of tasks. The table is from the target-agent’s perspective at the end of the task for the ‘*target-object*’, i.e. the object on which the task has to be performed. For example the task of ‘*making accessible*’ has been understood by the robot that after the task the object should be easier to be seen and reached by the ‘*target-agent*’ as well as it should be graspable and visibility of the object should be increased. Our robot is further equipped with the geometric interpretation of each symbolic descriptor. This task understanding will be used by various blocks of the proactive planner presented in this paper, as shown in fig. 3.

We prefer to use our robot’s capability to understand a set of basic tasks, however, in the absence of such capability or such capability being insufficient for a task, one can also provide a pre-scripted meaning to a task.

TABLE III
TASK UNDERSTANDING FROM MULTI-STATE VISUO-SPATIAL PERSPECTIVE FOR THE TARGET AGENT

Task	To See	To Reach	To Grasp	% Visibility
Show	Easier	Not Relevant	Not Relevant	Increase
Give	Easier	Easier	Graspable	Increase
Hide	Difficult	Not Relevant	Not Relevant	Decrease
Make Accessible	Easier	Easier	Graspable	Increase
Put Away	Maintain	Difficult	Difficult	Decrease
Hide Away	Difficult	Difficult	Difficult	Difficult



Fig. 2. Taxonomy of reach actions:(a) arm-shoulder reach, (b) arm-torso reach, (c) standing reach.

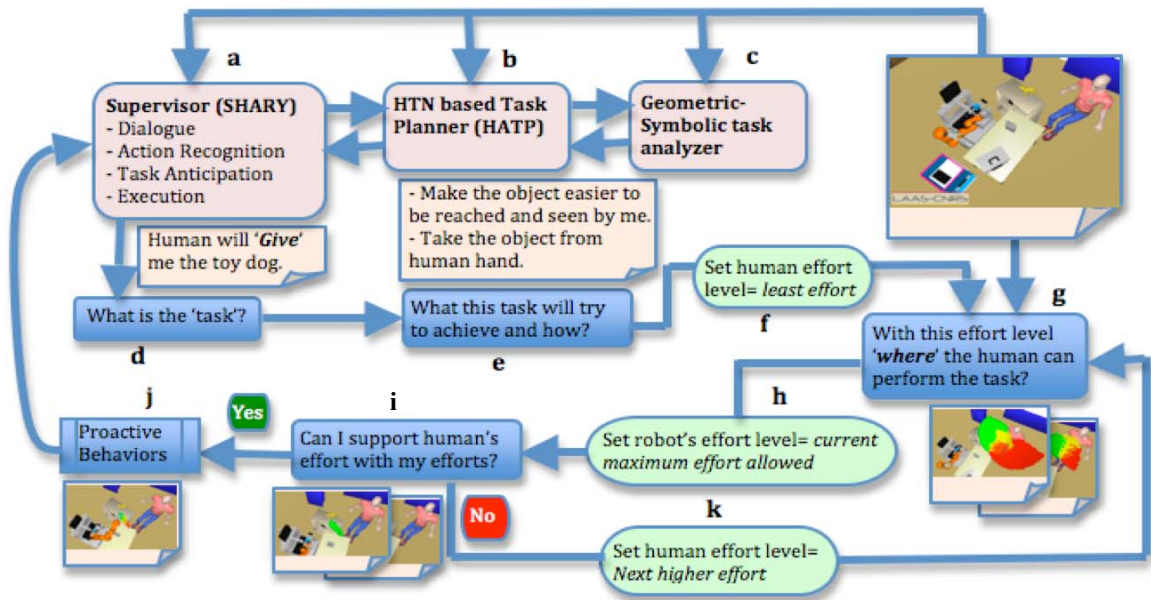


Fig. 3: Control flow of the overall system. Proactive planner takes the task requests from supervisor and returns appropriate proactive behaviors to it.

D. Proactive decision making

Fig. 3 shows the overview of the overall reasoning process for proactive decision-making. We will explain it with an example. As shown in fig. 3, SHARY is the robot supervisor, [27], which monitors the environment, decides and generates appropriate planning/re-planning and controls requests and executes them appropriately. Symbolic level plans for a particular task are generated by HATP, [19], a HTN based task planner, with two-way handshaking with its geometric counterpart, [26].

The example scenario shown on top-right of the fig. 3 is for a co-operative task of ‘clean the table’ of the CHRIS project [28]. The goal is to put all the objects laying on the tables in the trash bin on the right of the robot. Some objects are reachable and visible by the robot and some are exclusively by human. HATP generates the shared plan to achieve the task. In the plan, one co-operative action generated is to ask human to give the toy dog to the robot, which is currently not reachable by the robot, so that robot can put it into to trash bin. As the robot ‘understands’ the semantics of the task that human will try to make the object’s reachability and visibility easy for it and also the way human is expected to perform it i.e. by giving it in the robot’s hand, the robot will try to proactively support the task, as will be explained below.

The robot will reason about ‘where’ the human can

perform the task by setting different level of efforts for human starting from the lowest, as shown in table II, as an attempt to reduce the human’s effort. As the object is not already in human’s hand so robot will skip the test for *no_effort_required* and will set human effort level in ‘f’ block of fig. 3 as *arm_effort* for reachability of human and *head_effort* for visibility of human. This means that robot predicts if the human will stretch out the arm and only turn the head around, which are the places where he can give an object to someone. Fig. 4(a) shows the candidate points for giving the object by human for this level of effort. Green, red and yellow points show giving possibilities by right, left and both hands respectively, for the current level of human’s effort. As robot does a similar multi-state visuo-spatial analysis for itself also, in block ‘h’ of fig. 3, robot sets its maximum level of effort for the current task. This level is decided by the HATP, based on task, mutual effort balancing, social constraints, human’s preferences etc., and could be reset dynamically if required. For the current example this is set as *arm_effort*, which means robot should not turn or move its base, it can only move its arm for achieving current sub-task of getting the object from human. With this level of effort for itself, robot tests the feasibility of achieving the task, block ‘i’ of fig. 3. Depending upon the task, it performs different set of feasibility tests. For the current task in which robot has to take an object from

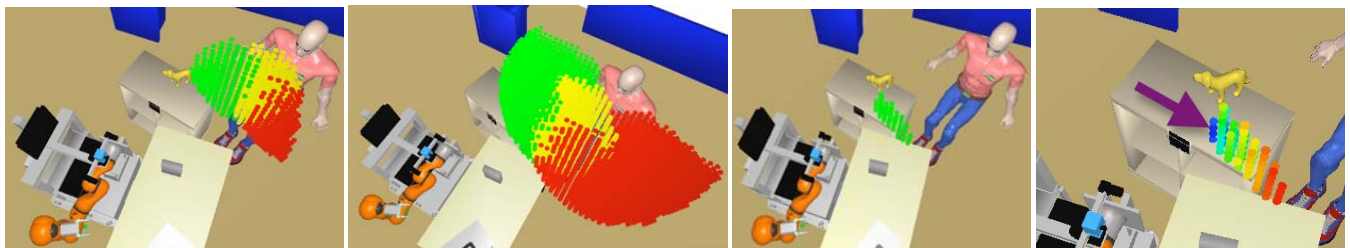


Fig. 4: Candidate points for giving an object by the human (green: by right hand, red: by left hand, yellow: by both hands) (a) from his current position, (b) if the human will make effort to move (lean forward or turn) his torso while remain seated. (c) Candidate points from where the robot can take the object for the effort level of (b) for the human, (d) weight assignment on the candidate points based on the closeness to the target-object, toy dog.

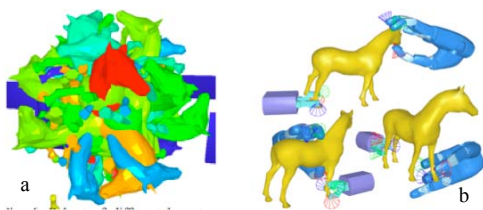


Fig. 5. (a) Different *placement* configurations of toy horse. (b) Sub-set of dual grasp by the human hand and the robot's gripper.

human, it performs feasibility test in the following way:

(i) Finds the points in the candidate points obtained for the human for his current effort level, which are also reachable and visible by robot with its maximum allowed effort.

(ii) Then it assigns weights to the resultant candidate points. Currently we assign weight based on the closeness to the target-object position, with the hypothesis that human need to put less effort in giving the object to the robot if he has to carry the object for less distance.

(iii) Then starting from the highest weighted candidate point it finds a possible collision free placement configurations of the object at that point. Fig. 5(a) shows all possible placement configurations of toy horse in free space.

(iv) Then for that placement of the object, it finds a grasp of the robot for that object. Since it is a task of handing over an object, it ensures that the selected grasp allows the human to grasp the object simultaneously. Fig. 5(b) shows a sub-set of dual grasps for the toy horse object.

In fact, all possible placement configurations independent of where to put as well as all possible grasps of an object are already known to the robot, for a collision free environment. This is a one-time computation for each new object the robot encounters in its 'lifetime', [22]. Then depending upon the environment, collision, visibility constraints, etc., robot just filters out the unwanted placements and grasps by putting the object at a particular point in a particular configuration.

(vi) Then the robot tries to find a collision free path for reaching to that point, [24].

If a particular candidate point passed all the feasibility tests for a particular task, then that particular point is considered to show the proactive behavior and the smooth trajectory is generated, [29], for execution.

For our current example, the robot fails to find the candidate points for test (i) itself as there was no point reachable by the robot with its current effort level for the points of fig 4(a). Hence the robot reached the block 'k' of fig. 3. Now it sets next effort level of human, which is

arm_torso_effort for reachability and *head_torso_effort* for visibility and loop back to block 'g'. Fig. 4(b) shows the candidate points for giving the object by human for this new level of effort, in which human can lean forward or turn around while remain seated. Then again by setting the same maximum allowed effort level for itself, the robot tests the possibility to support human's effort. This time robot finds a set of candidate points from where the robot can take an object from the human, green point cloud in fig. 4(c). Then as mentioned previously in point (ii), the resultant candidate points after the weight assignment have been shown in fig. 4(d). Blue point, having highest weight, will be preferred over the red points, having lowest weights. The robot obtains winner-feasible point, pointed in fig. 4(d), i.e. the first highest weighted candidate point, which has passed rest of the feasibility tests, and the algorithm reaches to block 'j'.

Block 'j', depending upon the task, returns appropriate data for showing proactive behaviors. For the current task it returns, the winner feasible point obtained in fig. 4(d), the corresponding levels of efforts for the human and for the robot, the trajectory to reach the point, the predicted end configuration of the robot for planned proactive behavior, as shown in fig. 6.

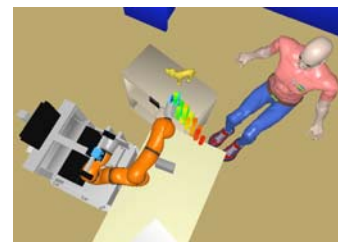


Fig. 6: Anticipated final configuration for the robot's proactive reach.

Depending upon the situation the supervisor will appropriately control the robot. In most of the situations the robot looks at the human and speaks "please give me the toy dog" and simultaneously moves the robot's arm on the trajectory towards the place to take. Supervisor actively monitors the environment during execution and if the human's intention or attention has changed or human is carrying the object away from the current feasible point, supervisor may decide to send appropriate re-planning requests either to symbolic or to proactive planner, depending upon the level of re-planning required. The object handing-over movement is identified by robot through its touch and force sensors associated with the gripper, which triggers to close the gripper for taking the object.

E. Adaptation for Make-Accessible task

Fig. 7 shows another scenario of clean the table demo.



Fig.7: Task of making an object accessible by the human to the robot. (a) Places on the support planes where the human can put the object, by right (green), both (yellow) and left (red) hands. (b) Candidate points where the robot can support human, blue has the highest weight requiring least effort from human. (c) Robot anticipates possible placement of the object on the box, from where it is feasible to take, while ensuring least effort from the human.



Fig.8: (a) Initial scenario for giving the object marked by red arrow. (b) The human is trying to give it to the robot by standing, *Whole Body Effort*, in the absence of proactive reach by the robot. In presence of proactive reach by robot’s arm (c) the human is giving it just by leaning forward, *Arm_Torso_Effort* and (d) giving it by stretching out his arm only, *Arm_Effort* in a different scenario, where the robot was able to support the *Arm_Effort* of the human.

Our HTN based task planner, HATP is capable of maintaining two streams of the plans: one for the human and another for the robot. Based on the temporal independency of sub-tasks at some particular point of execution, robot supervisor could decide to execute a part of plan in parallel by both: the human and the robot. For this it will generate a co-operative request like “Can you please make the grey-tape accessible to me, meanwhile I will put black tape in the trash bin?” In the absence of any proactive feedback from the robot, human can take and put the grey tape anywhere on the table, perhaps thinking about where to put so that robot can take. But if supervisor consults the proactive planner to know ‘where’ the human can put the grey tape with least effort, it generates the co-operative request as “Can you please make the grey-tape accessible to me, you can *put it on the box*, meanwhile I will put black tape in the trash bin?” Hence human is better guided about where to put.

The verbal proactive behavior mentioned above is achieved by using the same framework of the proactive planner of fig. 3, where instead of finding ‘where’ the human can hold the object, it finds ‘where’ the human can put the object for the robot to take it. Our robot is able to find the horizontal surfaces in the environment as the candidate points to place some object. So, it found that the human could put some object on the top of the toolbox, the candidate points have been shown in fig. 7(a). Fig. 7(b) shows the candidate points at step ‘i’ of the algorithm of fig 3. In this case robot modifies the feasibility tests of ‘give’ object described in section II-D. Now since human is not required to hand-over the object, robot will not perform the test of simultaneous dual grasp. Furthermore since now the placement on a support is required; robot will only consider the stable placement configurations on the corresponding support plane. We have dedicated modules to filters out the unstable placements configurations at a point on a support plane, based on the shape and size of object and of the support plane. Furthermore our robot is equipped with inter-

object spatial relation reasoning capability, such as ‘on’, ‘in’, ‘covered by’, ‘supported by’ etc., hence it can symbolically ground the candidate placement at a point of the support plane with ‘on object’. Fig. 7(c) shows the anticipated configuration of the robot to take the object and the possible placement of the object by the human for the candidate point of highest weight, *on the box*.

This is another aspect of proactive behavior: verbal proactive behavior, in which robot is trying to reduce the uncertainty in the human’s behavior as well as the human’s effort by verbally guiding where to perform the task by anticipating ‘where’ he can perform the task with least effort and the robot could support with its desired effort level.

III. EXPERIMENTAL RESULTS AND ANALYSIS

We have tested our system on different robots: JIDO and PR2. Robot uses Move3D, an integrated planning and visualization platform and through its various sensors maintains and updates the 3D world state. For object identification and localization it uses tags based stereovision system. For localizing human it uses data from *Kinect* motion sensor mounted on it.

Fig. 8(a) and (b), show the initial and final scenarios for the task of giving an object by human without proactive behavior. Human is standing to give the object, hence putting the *Whole Body Effort*. Fig 8(c) shows same task but robot is proactively reaching to take the object, hence reducing the human effort to *Arm_Torso_Effort* and *Arm_Effort* in fig. 8(d) having different spatial arrangement.

Fig. 9(a), shows another scenario for giving the object to the robot. One interesting observation is that without any proactive reaching behavior this particular user is holding the object and waiting for the robot to take. Whereas, as shown in fig. 9(b), in the presence of proactive behavior reaching behavior of the robot human is putting some effort to lean and give the object to the robot. This validates the



Fig.9: Task of giving an object to robot (a) In the absence of any proactive behavior this particular user is holding the object and waiting for the robot to take. (b) With proactive reach user is putting effort to give.



Fig.10: (a) Initial scenario for giving an object by the human to the PR2 robot. (b) The human is putting more effort in the absence of any proactive reach towards object behavior by robot. (c) The human is giving with less effort when robot is reaching out the object proactively.

studies of human-behavioral psychology that goal anticipation during action observation is influenced by synonymous action capabilities [6].

We have also tested our system on another robot PR2. Fig. 10(a), shows initial scenario. Fig. 10(b) shows user giving the object with *Arm_Torso_Effort* in the absence of robot’s proactive reach behavior, whereas fig. 10(c) shows human is giving the object with less effort, *Arm_Effort* to the robot when it has proactively moved its arm.

Fig.11: Human makes an object accessible to the robot. (a) In absence of verbal proactive behavior about where to put, human is putting with more effort, *Whole_Body_Effort*. (b) Human is putting with less effort *Arm_Torso_Effort*, when robot is proactively saying to put on the box. Red arrows show the initial positions of the object to make accessible.



Fig. 11 shows the task of making an object accessible by the human to the robot. When the robot does not behave proactively, human is putting the object closer to the robot with *Whole_Body_Effort*, Fig. 11(a). In the case of proactive behavior, the robot calculated the least effort places by the human to put, from where it can take the object and proactively says to put on the white box. In this case the human puts with less effort, *Arm_Torso_Effort*, fig. 11(b).

IV. DISCOVERIES BY PRELIMINARY USER STUDIES

We have performed preliminary user studies to validate as well as to refine the system. We will briefly discuss the user study of the task of giving some object by the human to the robot. We setup different scenarios having different relative positions of the robot, human and object. Broadly the scenario could be divided into two categories: (i) Human is sitting away from the robot and there are some furniture between them, similar to fig. 8(a). (ii) Human is sitting relatively close to the robot and there is no furniture between them, similar to fig. 9(a).

Each user has been exposed to two different behaviors of the robot, *NPB (Non Proactive Behavior)*: Robot just asks to the user “Please give me the <object name>” and waits in its current state, *PB (Proactive Behavior)*: robot asks the same but also starts moving its arm along the trajectory returned by the proactive planner. The order to exhibit PB or NPB to a particular user was random. The users were of three different categories based on their exposure to the human interactive robots: no, little or rich exposure. Users were told that robot will interact with them and they are expected to behave in their ways. They were not informed about the robot’s behaviors. There were a total of 12 participants.

After being demonstrated to both behaviors, each user was requested to fill a questionnaire with reference to first behavior *B1* and the second behavior *B2*. Note that for some users *B1* was *NPB (Non Proactive Behavior)* and for some *B1* was *PB (Proactive Behavior)*. Table IV shows answer to

one of the questions about the types of confusions the robot behavior has caused to the user. Table V summarizes the main observations from the user studies in terms of *NPB* and *PB*. The overall response was that with proactive behavior human was in less confusing states and also the human effort compared to non-proactive behavior has been reduced. Also users have reported that the robot seems to be more aware about the users’ capabilities in the cases it behaved proactively.

Note that the sums of % in these tables are not 100 as the users were allowed to mark multiple options or none.

Apart from the observations from the direct responses from users, we found the following interesting observations:

- For the cases where proactive behavior of robot has been demonstrated first, users seem to be biased towards expecting similar behavior for the later demonstration in which non-proactive behavior has been demonstrated. In such cases users’ responses were: “I thought that experiment had failed, since the robot didn’t move”, “I was waiting for the robot to take it from me.”

- For the cases where non-proactive behavior has been shown first, few users have been found ‘searching’ for the object to give, if the table-top environment was somewhat cluttered, even if the robot has asked to give the object by name. This suggests that such proactive behaviors also help in fetching the human’s attention to the object of interest.

But further user studies are absolutely required to validate these hypotheses.

V. CONCLUSION AND FUTURE WORKS

In this paper we have enabled our robot with an complementary issue of analyzing ‘where’ in order to behave proactively. This is achieved through multi-state

TABLE IV
USERS’ RESPONSES ABOUT THE TYPE OF CONFUSION

The confusion was: should I ... ->	go and give it to robot		stand up and give it to robot		put it somewhere for the robot to take		hold it somewhere and wait for the robot to move and take		wait for robot to show some activity	
	B1	B2	B1	B2	B1	B2	B1	B2	B1	B2
When B1 was proactive	0 %	12 %	0 %	0 %	0 %	25 %	0 %	37 %	0 %	25 %
When B2 was proactive	25 %	0 %	25 %	0 %	25 %	0 %	12 %	0 %	50 %	0 %

TABLE V
SUMMARY OF THE USERS’ EXPERIENCE ABOUT BOTH BEHAVIORS

User was confused about	Where to give	When to give	How to give
For Non Proactive Behavior	62 %	49 %	37 %
For Proactive Behavior	0 %	12 %	0 %

Overall appearance of robot	Human Effort Reducing	Demanding more human effort	Can’t say
In PB compared to NPB was	62 %	0 %	12 %
In NPB compared to PB was	0 %	50 %	12 %

Robot was reported to be ...	
more aware about human capabilities and possible confusion compared to PB with NPB	75 %
more supportive to the task compared to PB with NPB	62 %

visuo-spatial perspective taking based reasoning on ‘where’ the human can perform the task with different level of efforts. These are the novelties of our paper. Our research is inspired by studies of human behavioral psychology, which suggests that prediction of ‘where’ is an important aspect for co-

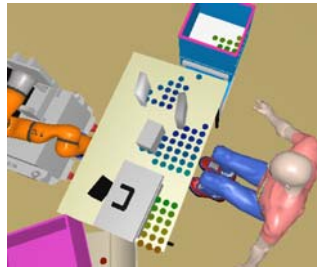


Fig. 12: Robot’s anticipation about the places where the human can hide some object from the robot, with a particular level of effort.

ordination. Our proactive planner generates verbal and non-verbal proactive behaviors. We have explored proactive behaviors for two tasks by the human: ‘give’ and ‘make accessible’. The same framework will be used to enable the robot with the proactive behavior for other tasks such as ‘show’, ‘hide’, ‘put away’, etc. The robot is able to anticipate where the human can perform these tasks with different level of efforts, as fig. 12 shows the candidate places where the human can hide some object from the robot, which enables the robot to involve in hide-and-seek like games. Also such capabilities will enable the robot to proactively decide to put some object, about which the human is oblivious, away from the human, to avoid to be hit by the human accidentally while doing something else.

Preminarily user studies show encouraging and supporting evidences to our intuition and hypotheses. But we feel the need of further user studies in same direction for establishing the observations as facts. Also we think that further user studies with a robot having two arms may reveal some entirely different confusing situations for the human user. Such as: if the robot does not show any proactive behavior, then which hand of the robot, the user has to select to give the object? Will there be any ambiguity? Will there be any improvement with proactive behavior, if yes then which types of proactive behaviors are required?

Such proactive behaviors are in fact essential building blocks of basic actions for complex socio-cognitive behaviors by incorporating expectation and intention. Hence validation and enhancement of such behaviors on long-term interaction with the users for a series of tasks are required.

REFERENCES

- [1] A. J. Schmid, O. Weede and H. Worn, “Proactive Robot Task Selection Given a Human Intention Estimate,” *IEEE RO-MAN*, 2007, pp.726-731.
- [2] T. Sakaki, “Estimation of intention of user arm motion for the proactive motion of upper extremity supporting robot,” *IEEE International Conference on Rehabilitation Robotics*, 2009.
- [3] W. Y. Kwon and I. H. Suh, “Towards proactive assistant robots for human assembly tasks,” *HRI*, 2011, pp. 175-176.
- [4] W. Y. Kwon and I. H. Suh, “Probabilistic Temporal Prediction for Proactive Action Selection,” In *IROS Workshop on Probabilistic Graphical Models in Robotics*, *GraphBot*, 2010.
- [5] O. C. Schrempf, U. D. Hanebeck, A. J. Schmid, H. Wörn, “A Novel Approach To Proactive Human-Robot Cooperation,” *IEEE RO-MAN*, 2005.
- [6] G. Gredebäck and O. Kochukhova, “Goal anticipation during action observation is influenced by synonymous action capabilities, a

- puzzling developmental study,” *Experimental Brain Research*, Springer Berlin/Heidelberg, vol. 202(2), pp. 493-497, December 2009.
- [7] G. Hoffman, “Anticipation in Human-Robot Interaction,” *AAAI Spring Symposium Series*, 2010.
- [8] N. Sebanz and G. Knoblich. Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, 1(2):353–367, 2009.
- [9] M. Tomasello, M. Carpenter, J. Call, T. Behne, and H. Moll, “Understanding and sharing intentions: The origins of cultural cognition,” *Behavioral and Brain Sciences*, 28, 675–735, 2005.
- [10] J. S. Jordan, and M. Hunsinger, “Learned patterns of action effect anticipation contribute to the spatial displacement of continuously moving stimuli,” *Journal of Experimental Psychology: Human Perception and Performance*, 34, 113–124, 2008.
- [11] Marcello L’Abbate, “Modeling Proactive Behavior of Conversational Interfaces,” Ph.D. dissertation, Fachbereich Informatik, Technische Universität Darmstadt, Darmstadt, Germany, 2007.
- [12] A. Cesta, G. Cortellessa, F. Pecora, and R. Rasconi, “Supporting interaction in the robocare intelligent assistive environment” *AAAI Spring Symposium on Interaction Challenges for Intelligent Assistants 2007*, pp. 18–25.
- [13] B. Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, vol. 12, p. 13-24, 2002.
- [14] J. Zwickel, “Agency attribution and visuo-spatial perspective taking,” *Psychonomic Bulletin*; Volume 16(6), Self, Other and Memory, Memory and Self-Understanding: Self-Concept - Self-Image - Self-Deception, December 2009, pp. 1089-1093.
- [15] H. Kockler, L. Scheef, R. Tepest, N. David, B. H. Bewernick, A. Newen, H.H. Schild, M. May, K. Vogetley, “Visuospatial perspective taking in a dynamic environment: Perceiving moving objects from a first-person-perspective induces a disposition to act,” *Consciousness and Cognition*, Volume 19(3), Self, Other and Memory, Memory and Self-Understanding: Self-Concept - Self-Image - Self-Deception, September 2010, pp. 690-701.
- [16] T. Taha, J. V. Mir’o, and G. Dissanayake, “Pomdp-based long-term user intention prediction for wheelchair navigation,” in *ICRA*, 2008, pp. 3920–3925.
- [17] T. Carlson and Y. Demirise, “Human-Wheelchair Collaboration Through Prediction of Intention and Adaptive Assistance,” in *ICRA*, 2008, pp 3926-3931.
- [18] T. V. Duong, H. H. Bui, D. Q. Phung, and S. Venkatesh, “Activity recognition and abnormality detection with the switching hidden semimarkov model,” in *CVPR* (1), 2005, pp. 838–845.
- [19] S. Alili, R. Alami and V. Montreuil, “A task planner for an autonomous social robot. DARS, 2008, Tsukuba, Japan.
- [20] A. K. Pandey and R. Alami, “Mightability Maps: A Perceptual Level Decisional Framework for Co-operative and Competitive Human-Robot Interaction,” *IROS*, 2010.
- [21] H. J. Choi and L. S. Mark, “Scaling affordances for human reach actions,” *Human Movement Science*, vol. 23, 2004, pp. 785-806.
- [22] J. P. Saut and D. Sidobre, “Efficient Models for Grasp Planning With A Multi-fingered Hand,” *Workshop on Grasp Planning and Task Learning by Imitation*, *IROS 2010*.
- [23] O. Khatib, E. Demircan, V. De Sapio, L. Sentis, T. Besier, and S. Delp. “Robotics-based Synthesis of Human Motion,” *Journal of Physiology Paris*. Vol. 103, no. 3-5, August 2009, pp. 211-219.
- [24] M. Gharbi, J. Cortés, T. Siméon, “A sampling-based path planner for dual-arm manipulation,” *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2008.
- [25] A. K. Pandey and R. Alami, “Towards Task Understanding through Multi-State Visuo-Spatial Perspective Taking for Human-Robot Interaction” *IJCAI 2011*, workshop on Agents Learning Interactively from Human Teachers, Barcelona, Spain.
- [26] S. Alili, A. K. Pandey, E. A. Sisbot and R. Alami, “Interleaving Symbolic and Geometric Reasoning for a Robotic Assistant”, *ICAPS Workshop on Combining Action and Motion Planning (CAMP 2010)*.
- [27] A. Clodic, H. Cao, S. Alili, V. Montreuil, R. Alami, and R. Chatila, “Shary: A supervision system adapted to human-robot interaction,” in *Experimental Robotics*, Springer Tracts in Advanced Robotics, 2009, vol. 54, pp. 229–238.
- [28] EU FP7 project CHRIS (Cooperative Human Robot Interaction Systems FP7 215805), <http://www.chrisfp7.eu/>
- [29] X. Broquere, D. Sidobre and K. Nguyen, “From Motion Planning to Trajectory Control with Bounded Jerk for Service Manipulator Robots,” *ICRA*, 2010, Anchorage, Alaska, USA.