What Is It to Implement a Human-Robot Joint Action?

Aurelie Clodic and Rachid Alami

Contents

Abstract

Joint action in the sphere of human–human interrelations may be a model for human–robot interactions. Human–human interrelations are only possible when several prerequisites are met, inter alia: (1) that each agent has a representation within itself of its distinction from the other so that their respective tasks can be coordinated; (2) each agent attends to the same object, is aware of that fact, and the two sets of "attentions" are causally connected; and (3) each agent understands the other's action as intentional. The authors explain how human–robot interaction can benefit from the same threefold pattern. In this context, two key problems emerge. First, how can a robot be programed to recognize its distinction from a human subject in the same space, to detect when a human agent is attending to something, to produce signals which exhibit their internal state and make decisions about the goal-directedness of the other's actions such that the appropriate predictions can be made? Second, what must humans learn about robots so they are able to interact reliably with them in view of a shared goal? This dual process is here examined by reference to the laboratory case of a human and a robot who team up in building a stack with four blocks.

Keywords

Human · Robot interaction · Joint action

Introduction

In this chapter, we present what is it to implement a joint action between a human and a robot. Joint action is "a social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment." (Sebanz et al. 2006: 70). We consider this implementation through a set of needed coordination processes to realize this joint action: Self-Other Distinction, Joint Attention, Understanding of Intentional Action, and Shared Task Representation. It is something that we have

A. Clodic (✉) · R. Alami
LAAS-CNRS, Université de Toulouse, CNRS, Artificial and Natural Intelligence Toulouse Institute (ANITI), Toulouse, France
e-mail: aurelie.clodic@laas.fr

already talked about in Clodic et al. (2017) but we will focus here on one example. Moreover, we will speak here about several elements that are components of a more global architecture described in Lemaignan et al. (2017). We introduce a simple human-robot collaborative to illustrate our approach. This example has been used as a benchmark in a series of workshop "toward a Framework for Joint Action" (fja.sciencesconf.org) and is illustrated in Fig. 1. A human and a robot have the common goal to build a stack with four blocks. They should stack the blocks in a specific order (1, 2, 3, 4). Each agent participates in the task by placing his/its blocks on the stack. The actions available to each agent are the following: take a block on the table, put a block on the stack, remove a block from the stack, place a block on the table, and give a block to the other agent.

This presentation is a partial point of view regarding what is and can be done to implement a joint action between a robot and a human since it presents only one example and a set of software developed in our lab. It only intends to explain what we claim is needed to enable a robot to run such a simple scenario.

At this point, it has to be noticed that from a philosophical point of view, we have been taught that some philosophers such as Seibt (2017) stressed that the robotics intentionalist vocabulary that we use is considered as problematic especially when robots are placed in social interaction spaces. In the following, we will use this intentionalist vocabulary in order to describe the functionalities of the robot, such as "believe" and "answers," because this is the way we describe our work in robotics and AI communities. However, to accommodate the philosophical concern, we would like to note that this can be considered as shorthand for "the robot simulates the belief," "the robot simulates an answer," etc. Thus, whenever robotic behavior is described with a verb that normally characterizes a human action, these passages can be read as a reference to the robot's simulation of the relevant action.
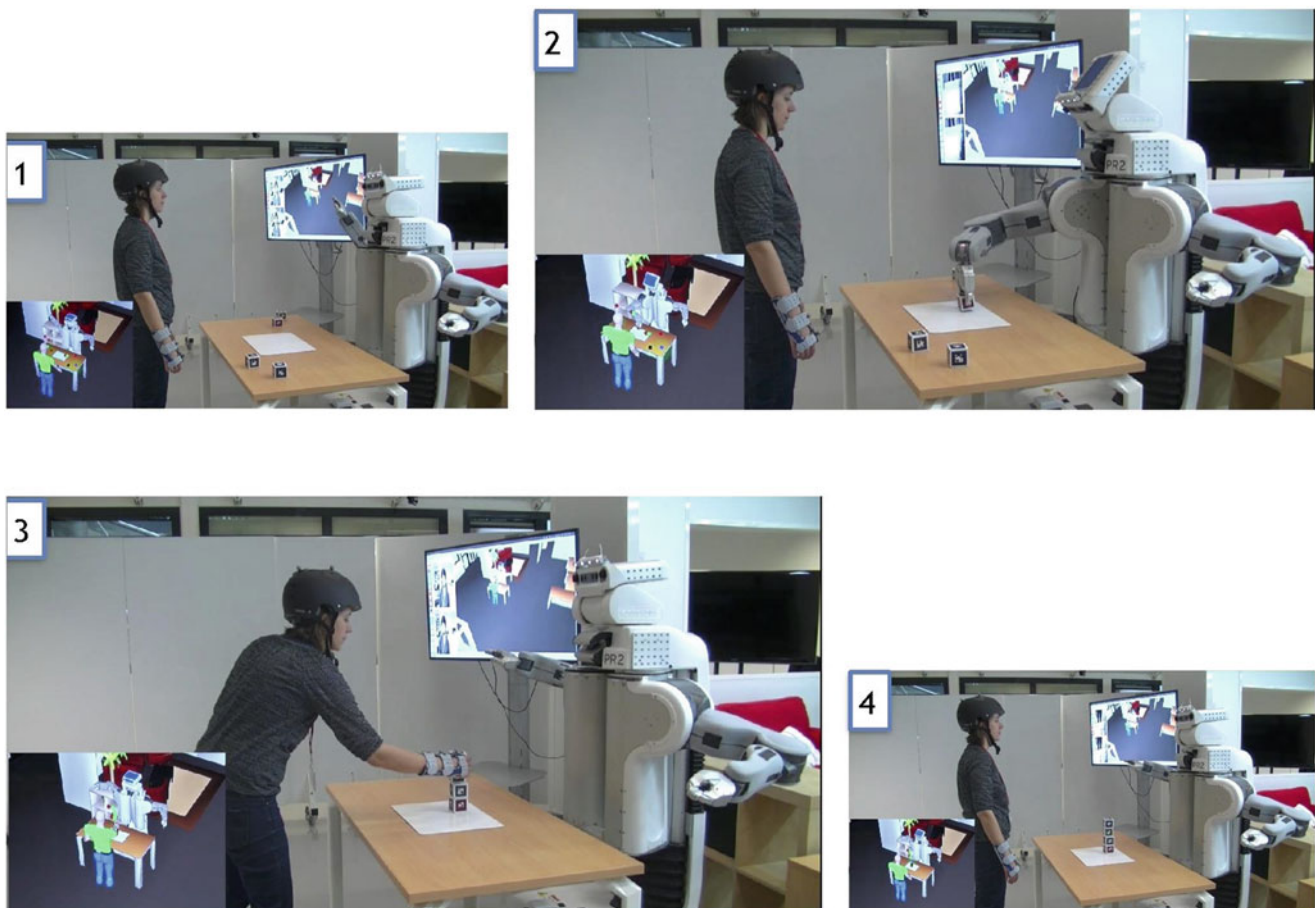
Fig. 1 A simple human–robot interaction scenario: A human and a robot have the common goal to build a stack with four blocks. They should stack the blocks in a specific order (1, 2, 3, 4). Each agent participates in the task by placing his/its blocks on the stack. The actions available to each agent are the following: take a block on the table, put a block on the stack, remove a block from the stack, place a block on the table, and give a block to the other agent. Also, the human and the robot observe one another. Copyright laas/cnrs https://homepages.laas.fr/aclodic

Self-Other Distinction

The first coordination process is Self-Other Distinction. It means that "for shared representations of actions and tasks to foster coordination rather than create confusion, it is important that agents also be able to keep apart representations of their own and other's actions and intentions" (Pacherie 2012: 359).

Regarding our example, it means that each agent should be able to create and maintain a representation of the world for its own but also from the point of view of the other agent. In the following, we will explain what the robot can do to build this kind of representation. The way a human (can) builds such representation for the robot agent (and on which basis) is still an open question.

Joint Attention

The second coordination process is Joint Attention. Attention is the mental activity by which we select among items in our perceptual field, focusing on some rather than others (see Watzl 2017). In a joint action setting, we have to deal with joint attention, which is more than the addition of two persons' attention. "The phenomenon of joint attention involves more than just two people attending to the same object or event. At least two additional conditions must be obtained. First, there must be some causal connection between the two subjects' acts of attending (causal coordination). Second, each subject must be aware, in some sense, of the object as an object that is present to both; in other words, the fact that both are attending to the same object or event should be open or mutually manifest (mutual manifestness)" (Pacherie 2012: 355).

On the robot side, it means that the robot must be able to detect and represent what is present in the joint action space, i.e., the joint attention space. It needs to be equipped with situation assessment capabilities (Lemaignan et al. 2018; Milliez et al. 2014).

In our example, illustrated in Fig. 2, it means that the robot needs to get:
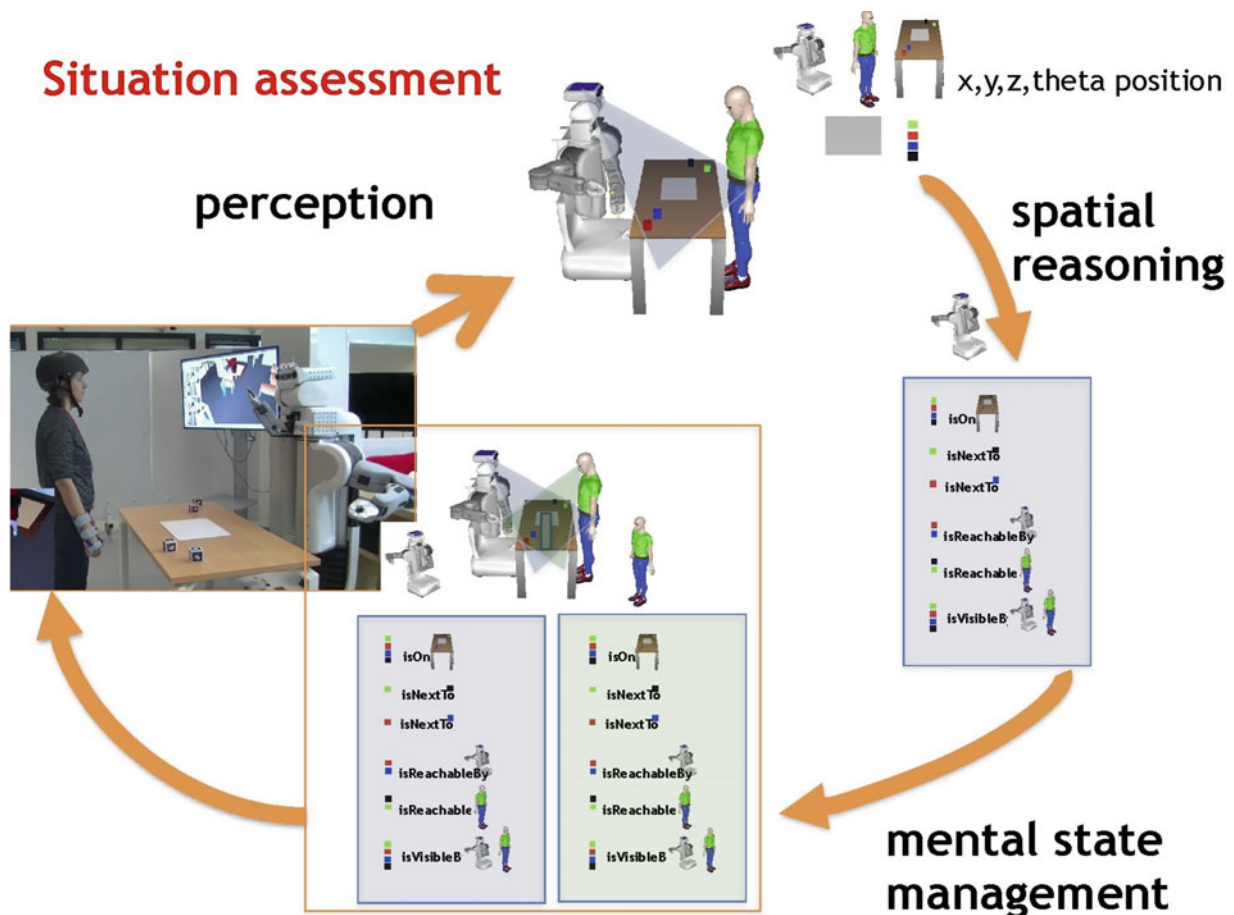
Fig. 2 Situation Assessment: the robot perceives its environment, builds a model of it, and computes facts through spatial reasoning to be able to share information with the human at a high level of abstraction and realizes mental state management to infer human knowledge. Copyright laas/cnrs https://homepages.laas.fr/aclodic

Fig. 3 What can we infer viewing this robot? There is no standard interface for the robot so it is difficult if not impossible to infer what this robot is able to do and what it is able to perceive (from its environment but also from the human it interacts with). Copyright laas/cnrs https://homepages.laas.fr/aclodic
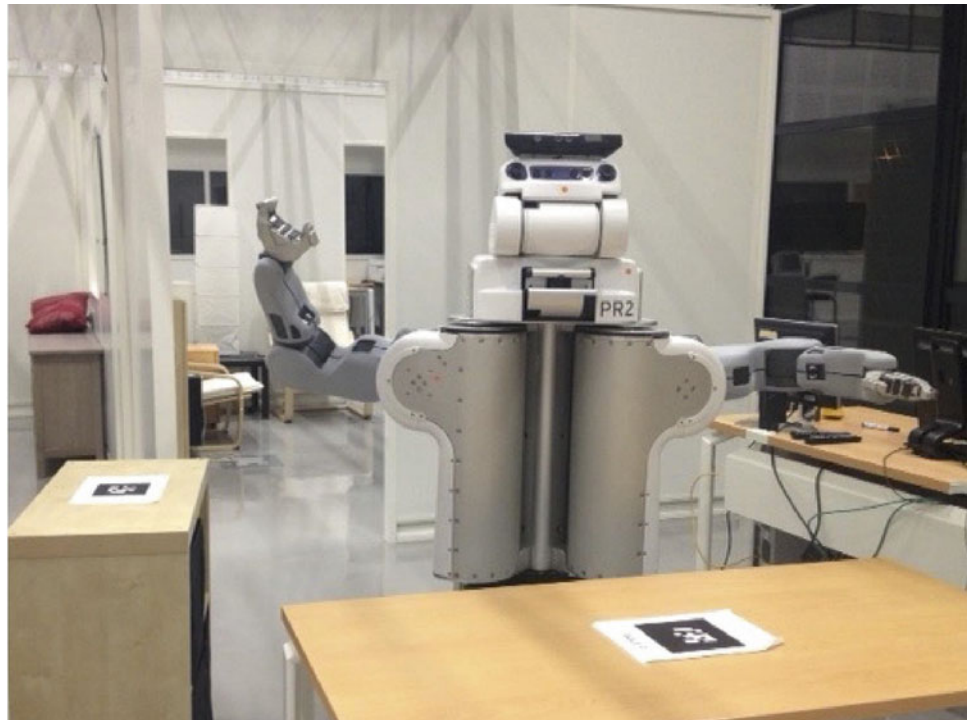
- its own position, that could be done for example by positioning the robot on a map and localizing it with the help of its laser (e.g., using amcl localization (http://wiki.ros.org/amcl) and gmapping (http://wiki.ros.org/gmapping))
- the position of the human with whom it interacts with (e.g., here it is tracked through the use of a motion capture system, that's why the human wears a helmet and a wrist brace. So more precisely, in this example, the robot has access to the head position and the right hand position)
- the position of the objects in the environment (e.g., here, a QR-code (https://en.wikipedia.org/wiki/QR_code) has been glued on each face of each block. These codes, and so, the blocks are tracked with one of the robot cameras. We get the 3D position of each block in the environment (e.g., with http://wiki.ros.org/ar_track_alvar))

However, each position computed by the robot is given as x, y, z, and theta position in a given frame. We cannot imagine to use such information to elaborate a verbal interaction with the human: "please take the block at position $x = 7.5$ m, $y = 3.0$ m, $Z = 1.0$ m, and theta $= 3.0$ radians in the frame map...". To overcome this limit, we must transform each position in an information that is understandable by (and hence shareable with) the human, e.g., *(RedBlock is On Table)*. We can also compute additional information such as *(GreenBlock is Visible By Human)* or *(BlueBlock is Reachable By Robot)*. This is what we call "spatial reasoning." Finally, the robot must also be aware that the information available to the human can be different from the one it has access to, e.g., an obstacle on the table can prevent her/him to see what is on the table. To infer the human knowledge, we compute all the information not only from the robot point of view but also from the human position point of view (Alami et al. 2011; Warnier et al. 2012; Milliez et al. 2014), it is what we call "mental state management."

On the human side, we can infer that the human is able to have the same set of information from the situation. But joint attention is more than that. We have to take into account "mutual manifestness," i.e., "(...) each subject must be aware in some sense, of the object as an object that is present to both; in other words the fact that both are attending to the same object or event should be open or mutually manifest..." (Pacherie 2012: 355). It raises several questions. How can a robot exhibit joint attention? What cues the robot should exhibit to let the human to infer that joint attention is met? How can a robot know that the human it interacts with is really involved in the joint task? What are the cues that should be collected by the robot to infer joint attention? These questions are still open questions. To answer them, we have to work particularly on the way to make the robot more understandable and more legible. For example, viewing this robot in Fig. 3, what can one infer about its capabilities?

Understanding of Intentional Action

"Understanding intentions is foundational because it provides the interpretive matrix for deciding precisely what it is that someone is doing in the first place. Thus, the exact same physical movement may be seen as giving an object,

sharing it, loaning it, moving it, getting rid of it, returning it, trading it, selling it, and on and on—depending on the goals and intentions of the actor" (Tomasello et al. 2005: 675). Understanding of intentional action could be seen as a building block of understanding intentions, it means that each agent should be able to read its partner's actions. To understand an intentional action, an agent should, when observing a partner's action or course of actions, be able to infer their partner's intention. Here, when we speak about partner's intention we mean its goal and its plan. It is linked to action-to-goal prediction (i.e., viewing and understanding the on-going action, you are able to infer the underlying goal) and goal-to-action prediction (i.e., knowing the goal you are able to infer what would be the action(s) needed to achieve it).

On the robot side, it means that it needs to be able to understand what the human is currently doing and to be able to predict the outcomes of the human's actions, e.g., it must be equipped with action recognition abilities. The difficulty here is to frame what should and can be recognized since the spectrum is vast regarding what the human is able to do. A way to do that is to choose to consider only a set of actions framed by a particular task.

On the other side, the human needs to be able to understand what the robot is doing, be able to infer the goal and to predict the outcomes of the robot's actions. It means, viewing a movement, the human should be able to infer what is the underlying action of the robot. That means the robot should perform movement that can be read by the human. Before doing a movement, the robot needs to compute it, it is motion planning. Motion planning takes as inputs an initial and a final configuration (for manipulation, it is the position of the arms; for navigation, it is the position of the robot basis). Motion planning computes a path or a trajectory from the initial configuration to the final configuration. This path could be possible but not understandable and/or legible and/or predictable for the human. For example, in Fig. 4, on the left, you see a path which is possible but should be avoided if possible, the one on the right should be preferred.

In addition, some paths could be also dangerous and/or not comfortable for the human, as illustrated in Fig. 5. Human-aware motion planning (Sisbot et al. 2007; Kruse et al. 2013; Khambhaita and Alami 2017a, b) has been developed to enable the robot to handle the choice of a path that is acceptable, predictable, and comfortable to the human the robot interacts with.

Figure 6 shows an implementation of a human-aware motion planning algorithm (Sisbot et al. 2007, 2010; Sisbot and Alami 2012) which takes into account safety, visibility, and comfort of the human. In addition, this algorithm is able to compute a path for both the robot and the human, which can solve a situation where a human action is needed or can be used to balance effort between the two agents.

However, it is not sufficient. When a robot is equipped with something that looks like a head, for example, people tend to consider that it should act like a head because people anthropomorphize. It means that we need to consider the entire body of the robot and not only the base or the arms of the robot for the movement even if it is not needed to achieve the action (e.g., Gharbi et al. 2015; Khambhaita et al. 2016). This could be linked to the concept of coordination smoother which is "any kind of modulation of one's movements that reliably has the effect of simplifying coordination" (Vesper et al. 2010, p. 1001).

Fig. 4 Two final positions of the arm of the robot to get the object. The one at right is better from an interaction point of view since it is easily understandable by the human. However, from a computational point of view (and even from an efficiency if we just consider the robot action that needs to be performed) they are equivalent. Consequently, we need to take these features explicitly into account when planning robot motions. That is what human-aware motion planning aims to achieve. Copyright laas/cnrs https://homepages.laas.fr/aclodic

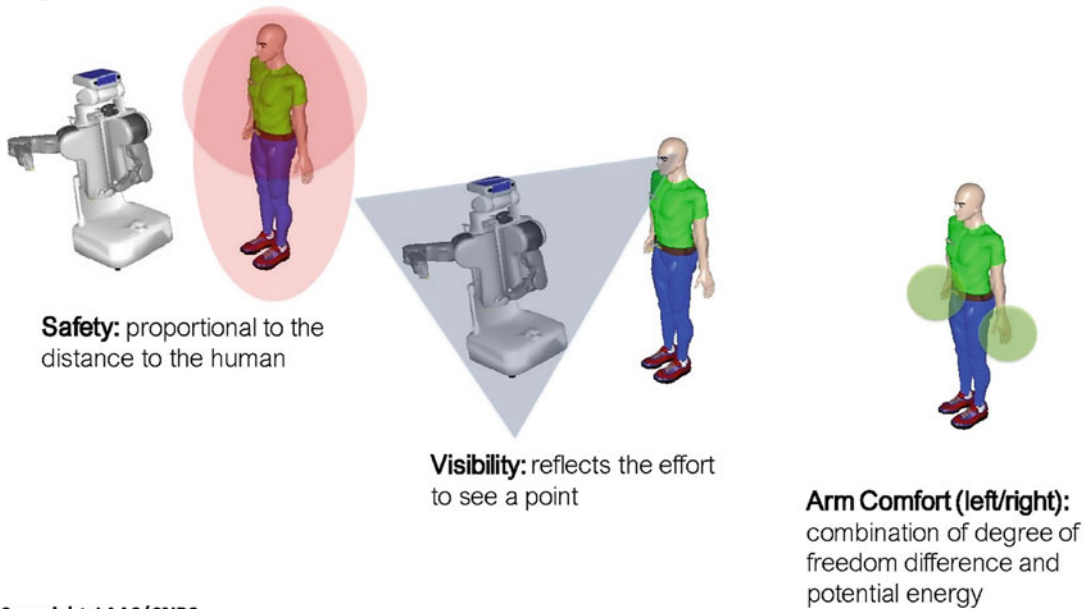**Human-Aware Motion Planning**

Copyright: LAAS/CNRS

Fig. 5 Not "human-aware" positions of the robot. Several criteria should be taken into account, such as safety, comfort, and visibility. This is for the hand-over position but also for the overall robot position itself. Copyright laas/cnrs https://homepages.laas.fr/aclodic

Fig. 6 An example of human-aware motion planning algorithm combining three criteria: safety of the human, visibility of the robot by the human, and comfort of the human. The three criteria can be weighed according to their importance with a given person, at a particular location or time of the task. Copyright laas/cnrs https://homepages.laas.fr/aclodic

Shared Task Representations

The last coordination process is shared task representations. As emphasized by Knoblich and colleagues (Knoblich et al. 2011), shared task representations play an important role in goal-directed coordination. Sharing representations can be considered as putting in perspective all the processes already described, e.g., knowing that the robot and the human track the same block in the interaction scene through joint attention and that the robot is currently moving this block in the direction of the stack by the help of intentional action understanding make sense in the context of the robot and the human building a stack together in the framework of a joint action.

To be able to share task representations, we need to have the same ones (or a way to understand them). We developed a Human-Aware Task Planner (HATP) based on Hierarchical Task Network (HTN) representation (Alami et al. 2006; Montreuil et al. 2007; Alili et al. 2009; Clodic et al. 2009; Lallement et al. 2014). The domain representation is illustrated in Fig. 7, it is composed of a set of actions (e.g., placeCube) and a set of tasks (e.g., buildStack) which combine action(s) and task(s). One of the advantages of such representation is that it is human readable. Here, placeCube

(Agent R, Cube C, Area A) means that for an Agent R, to place the Cube C in the Area A, the precondition is that R has in hand the Cube C and the effects of the action is that R has no more the Cube C in hand but the object C is on the stack of Area A. It is possible to add cost and duration to each action if we want to weigh the influence of each of the actions.

On the other hand, BuildStack is done by adding a cube (addCube) and then continue to build the stack (buildStack). Then each task is also refined until we get an action. HATP computes a plan both for the robot and the human (or humans) it interacts with as illustrated in Fig. 8. The workload could be balanced between the robot and the human; moreover, the system enables to postpone the choice of the actor at execution time (Devin et al. 2018). However, one of the drawbacks of such representation is that it is not expandable. Once the domain is written, you cannot modify it. One idea could be to use reinforcement learning. However, reinforcement learning is difficult to use "as is" in a human–robot interaction case. The reinforcement learning system needs to test any combination of actions to be able to learn the best one which could lead to nonsense behavior of the robot. This can be difficult to interpret for the human it interacts with and it will be difficult for him to interact with the robot,
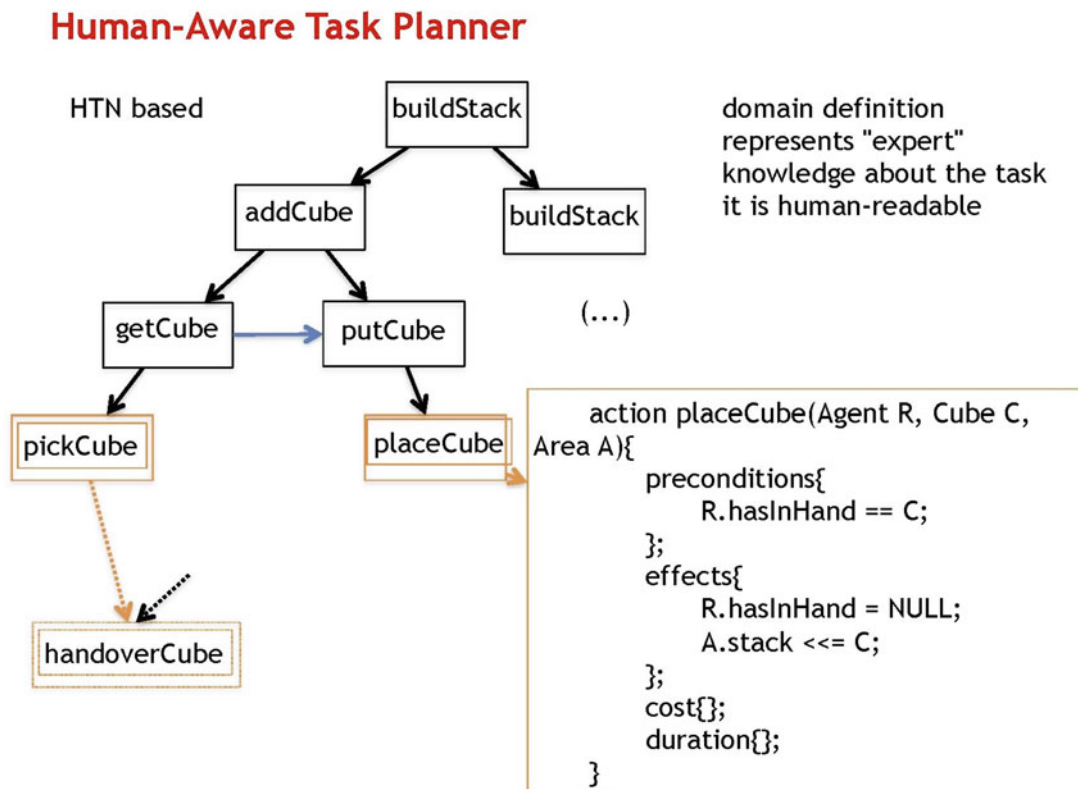
Fig. 7 HATP domain definition for the joint task buildStack and definition of the action placeCube: The action placeCube for an Agent R, a Cube C in an Area A, could be defined as follows. The precondition is that Agent R has the Cube C in hand before the action, the effect of the action is that Agent R does not have the Cube C anymore and the cube C is on the stack in Area A. Task buildStack combines addCube and buildStack. Task addCube combines getCube and putCube. Task getCube could be done either by picking the Cube or doing a handover. Copyright laas/cnrs https://homepages.laas.fr/aclodic
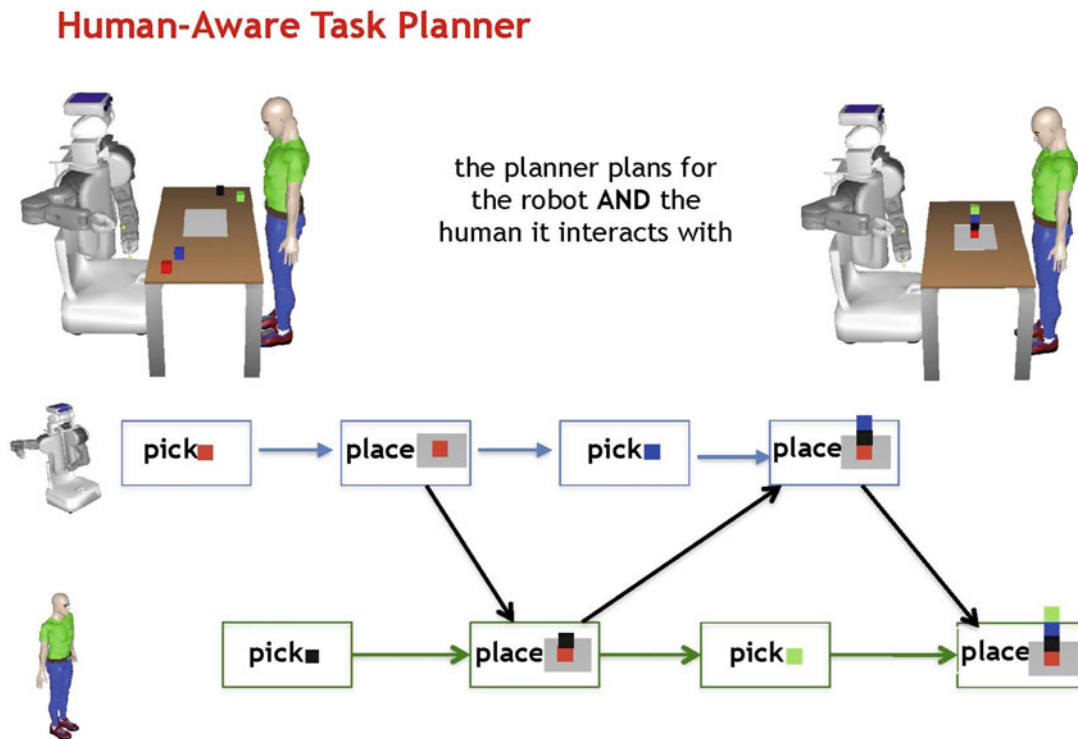
Fig. 8 HATP shared (human and robot) plan example for the stack of cubes example. Copyright laas/cnrs https://homepages.laas.fr/aclodic

and will lead to learning failure. To overcome this limitation, we have proposed to mix the two approaches by using HATP as a bootstrap for a reinforcement learning system (Renaudo et al. 2015; Chatila et al. 2018).

With a planning system as HATP, we have a plan for both the robot and the human it interacts with but this is not enough. If we follow Knoblich and colleagues (Knoblich et al. 2011) idea, shared task representations do not only specify in advance what the respective tasks of each of the coagents are, they also provide control structures that allow agents to monitor and predict what their partners are doing, thus enabling interpersonal coordination in real time. This means that the robot not only need the plan, but also ways to monitor this plan. Besides the world state (cf. Fig. 2 section regarding situation assessment) and the plan, we developed a monitoring system that enables the robot to infer plan status and action status both from its point of view and from the point of view of the human as illustrated Fig. 9 (Devin and Alami 2016; Devin et al. 2017). With this information, the robot is able to adapt its execution in real time. For example, there may be a mismatch between action status on the robot side and on the human side (e.g., the robot waiting for an action from the human). Equipped with this monitoring, the robot can detect the issue and warn. The issue can be at plan status level, e.g., the robot considering that the plan is no longer achievable while it detects that the human continues to act.

Conclusion

We have presented four coordination processes needed to realize a joint action. Taking these different processes into account requires the implementation of dedicated software: self-other distinction → mental state management; joint attention → situation assessment; understanding of intentional action → action recognition abilities as well as human-aware action (motion) planning and execution; shared task representations → human-aware task planning and execution as well as monitoring.

The execution of a joint action requires not only for the robot to be able to achieve its part of the task but to achieve it in a way that is understandable to the human it interacts with and to take into account the reaction of the human if any. Mixing execution and monitoring requires making some choices at some point, e.g., if the camera is needed to do an action, the robot cannot use it to monitor the human if it is not in the same field of view. These choices are made by the supervision system which manages the overall task execution from task planning to low-level action execution.

We talked a little bit about how the human was managing these different coordination processes in a human–robot interaction framework and about the fact that there was still some uncertainty about how he was managing things. We believe that it may be necessary in the long term to give the human the means to better understand the robot at first.

Shared Plan Management

Taking others mental state during execution
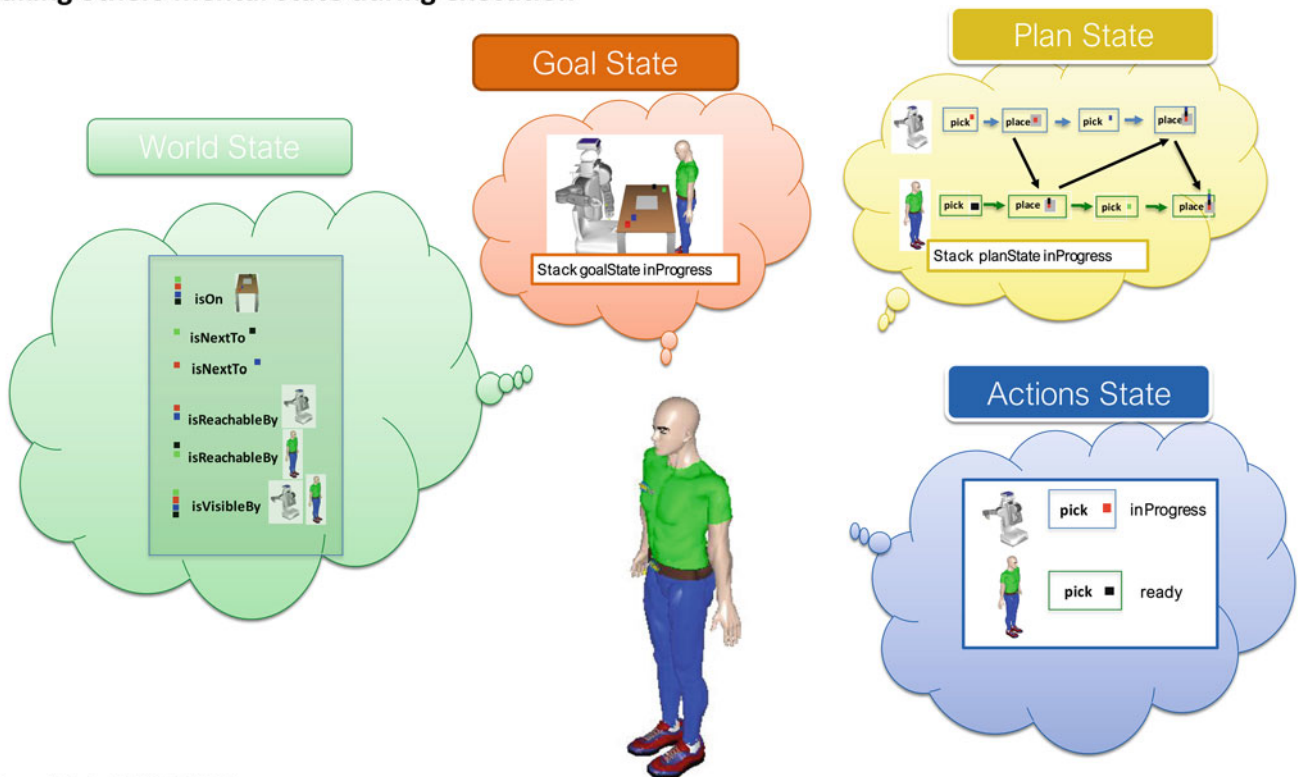
Copyright: LAAS/CNRS

Fig. 9 Monitoring the human side of the plan execution: besides the world state, the robot computes the state of the goals that need to be achieved, the status of the on-going plans as well of each action. It is done not only from its point of view but also from the point of view of the human. Copyright laas/cnrs https://homepages.laas.fr/aclodic

Finally, what has been presented in this chapter is partial for at least two reasons. First, we have chosen to present only work done in our lab but this work already covers the execution of an entire task and in an interesting variety of dimensions. Second, we make the choice to not mention the way to handle communication or dialog, to handle data management or memory, to handle negotiation or commitments management, to enable learning, to take into account social aspects (incl. privacy) or even emotional ones, etc. However, it gives a first intuition to understand what needs to be taken into account to make a human–robot interaction successful (even for a very simple task).

References

Alami, R., Clodic, A., Montreuil, V., Sisbot, E. A., & Chatila, R. (2006). Toward human-aware robot task planning. In *AAAI spring symposium: To boldly go where no human-robot team has gone before* (pp. 39–46). Palo Alto, CA.

Alami, R, Warnier, M., Guitton, J., Lemaignan, S., & Sisbo, E. A. (2011). When the robot considers the human. In *15th Int. Symposium on Robotics Research*.

Alili, S., Alami, R., & Montreuil, V. (2009). A task planner for an autonomous social robot. In H. Asama, H. Kurokawa, J. Ota, & K. Sekiyama (Eds.), *Distributed autonomous robotic systems* (Vol. 8, pp. 335–344). Berlin/Heidelberg: Springer.

Chatila, R., Renaudo, E., Andries, M., Chavez-Garcia, R. O., Luce-Vayrac, P., Gottstein, R., Alami, R., Clodic, A., Devin, S., Girard, B., & Khamassi, M. (2018). Toward self-aware robots. *Frontiers in Robotics and AI*, 51–20. https://doi.org/10.3389/frobt.2018.00088.

Clodic, A., Cao, H., Alili, S., Montreuil, V., Alami, R., & Chatila, R. (2009). Shary: A supervision system adapted to human-robot interaction. In R. Clodic, H. Cao, S. Alili, V. Montreuil, R. Alami, & R. Chatila (Eds.), *Experimental robotics* (pp. 229–238). Berlin/Heidelberg: Springer.

Clodic, A., Pacherie, E., Alami, R., & Chatila, R. (2017). Key elements for human-robot joint action. In R. Hakli & J. Seibt (Eds.), *Sociality and normativity for robots* (pp. 159–177). Cham: Springer International Publishing.

Devin, S., & Alami, R. (2016). An implemented theory of mind to improve human-robot shared plans execution. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI)* (pp. 319–326). Christchurch, New Zealand, March 7–10, 2016.201, IEEE Press.

Devin, S., Clodic, A., & Alami, R. (2017). About decisions during human-robot shared plan achievement: Who should act and how? In A. Kheddar, E. Yoshida, S. S. Ge, K. Suzuki, J. J. Cabibihan, F. Eyssel, & H. He (Eds.), *Social robotics* (Vol. 10652, pp. 453–463). Cham: Springer International Publishing.

Devin, S., Vrignaud, C., Belhassein, K., Clodic, A., Carreras, O., & Alami, R. (2018). Evaluating the pertinence of robot decisions in a human-robot joint action context: The PeRDITA questionnaire. In *27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) 2018* (pp. 144–151), IEEE Press.

Gharbi, M., Paubel, P. V., Clodic, A., Carreras, O., Alami, R., & Cellier, J. M. (2015). Toward a better understanding of the communication cues involved in a human-robot object transfer. In *24th and Human Interactive Communication (RO-MAN), 2015* (pp. 319–324). IEEE Press.

Khambhaita, H., & Alami, R. (2017a). Assessing the social criteria for human-robot collaborative navigation: A comparison of human-aware navigation planners. In *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) 2017* (pp. 1140–1145), IEEE Press.

Khambhaita, H., & Alami, R. (2017b). *Viewing robot navigation in human environment as a cooperative activity*. Paper presented at the international symposium on robotics research (ISSR), Puerto Varas, Chile, 11–14 December 2017.

Khambhaita, H., Rios-Martinez, J., & Alami, R. (2016). *Head-body motion coordination for human aware robot navigation*. Paper presented at the 9th International Workshop on Human-Friendly Robotics (HFR 2016), Genoa, Italy, 29-30 September 2016.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 54, pp. 59–101). San Diego: Academic, Elsevier.

Kruse, T., Pandey, A. K., Alami, R., & Kirsch, A. (2013). Human-aware robot navigation: A survey. *Robotics and Autonomous Systems, 61*(12), 1726–1743. https://doi.org/10.1016/j.robot.2013.05.007.

Lallement, R., De Silva, L., & Alami, R. (2014). *HATP: An HTN planner for robotics*. Paper presented at the 2nd ICAPS Workshop on Planning and Robotics, Portsmouth, 22–23 June 2014.

Lemaignan, S., Warnier, M., Sisbot, E. A., Clodic, A., & Alami, R. (2017). Artificial cognition for social human–robot interaction: An implementation. *Artificial Intelligence, 247*, 45–69. https://doi.org/10.1016/j.artint.2016.07.002.

Lemaignan, S., Sallami, Y., Wallbridge, C., Clodic, A., Belpaeme, T., & Alami, R. (2018). *UNDERWORLDS: Cascading situation assessment for robots*. Paper presented at the IEEE/RSJ international conference on intelligent robots and systems (IROS 2018), Madrid, Spain, 1–5 October 2018.

Milliez, G., Warnier, M., Clodic, A., & Alami, R. (2014). A framework for endowing an interactive robot with reasoning capabilities about perspective-taking and belief management. In *The 23rd IEEE inter-national symposium on robot and human interactive communication* (pp. 1103–1109). IEEE Press.

Montreuil, V., Clodic, A., & Alami, R. (2007). *Planning human centered robot activities*. Paper presented on the 2007 IEEE international conference on systems, man and cybernetics, Montreal, Canada, 7–10 October 2007.

Pacherie, E. (2012). The phenomenology of joint action: Self-agency versus joint agency. In A. Seeman (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 343–389). Cambridge: MIT Press.

Renaudo, E., Devin, S., Girard, B., Chatila Alami, R., Khamassi, M., & Clodic, A. (2015). *Learning to interact with humans using goal-directed and habitual behaviors*. Paper presented at the RoMan 2015 Workshop on Learning for Human-Robot Collaboration, 31 August 2015.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences, 10*, 70–76. https://doi.org/10.1016/j.tics.2005.12.009.

Seibt, J. (2017). Towards an ontology of simulated social interaction: Varieties of the "as if" for robots and humans. In R. Hakli & J. Seibt (Eds.), *Sociality and normativity for robots* (pp. 11–39). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-53133-5_2.

Sisbot, E. A., & Alami, R. (2012). A human-aware manipulation planner. *IEEE Transactions on Robotics, 28*(5), 1045–1057.

Sisbot, E. A., Marin-Urias, L. F., Alami, R., & Siméon, T. (2007). A human aware mobile robot motion planner. *IEEE Transactions on Robotics, 23*(5), 874–883.

Sisbot, E. A., Marin-Urias, L., Broquère, X., Sidobre, D., & Alami, R. (2010). Synthesizing robot motions adapted to human presence. *International Journal of Social Robotics, 2*(3), 329–343.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences, 28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks, 23*(8–9), 998–1003. https://doi.org/10.1016/j.neunet.2010.06.002.

Warnier, M., Guitton, J., Lemaignan, S., & Alami, R. (2012). When the robot puts itself in your shoes. Explicit geometric management of position beliefs. In *The 21st IEEE international symposium on robot and human interactive communication (RO-MAN) 2012* (pp. 948–954), IEEE Press.

Watzl, S. (2017). *Structuring mind: The nature of attention and how it shapes consciousness*. Oxford: Oxford University Press.