IAC-10.B3.6.-A5.3.7

THE ROLE OF NATURAL INTERACTION IN ASTRONAUT-ROBOT COOPERATION

S. Heikkilä
Aalto University, Helsinki, Finland, Seppo.Heikkila@tkk.fi

This paper reviews five different methods to enable natural astronaut-robot interaction in future planetary exploration missions. Natural astronaut-robot interaction is used here to refer to interaction that is inspired by human-human interaction. All the reviewed methods are presented in the form that they could be directly implemented as external user interface modules to provide additional flexibility to the astronaut-robot communication. These natural communication methods mainly provide only additional value to the astronaut-robot cooperation because they do not replace any existing communication methods but instead introduce new ones that astronauts are already accustomed to use in their daily lives on Earth.

I. INTRODUCTION

The next manned missions to the surfaces of the Moon and Mars will be longer and more complex than any of the previous human spaceflights. The number of tasks and devices that astronauts need to perform and operate, often without any assistance from the ground control, is expected to increase significantly. One way to cope with this increased complexity is to develop interfaces that can support the human cognitive processes, i.e. develop interfaces that are based on the way people naturally interact and process information.

Natural interaction is used here to refer to interaction that is inspired by the human-human form of interaction. This type of natural human-robot interaction has been frequently envisioned in both science fiction and space robotics research as the way we humans would communicate with robots in the future [1, 2].

This paper describes research done in the SpacePartner project [3], the objective of which is to develop methods to enable natural and seamless astronaut-robot interaction. The test platform used in the project is a centaur-type service robot, called WorkPartner, which was initially developed at Aalto University to assist humans with light outdoor tasks. The WorkPartner robot, shown in Fig. 1, is described in detail in [4, 5].

The paper structure is as follows. Section II begins by discussing about the current status of human-robot interaction development, the problems that would need to be solved and why natural human-robot interaction could be part of the solution. Then in Section III four existing natural human-robot interaction methods are presented and analysed from the astronaut-robot planetary exploration application point of view. Section IV introduces idea and preliminary test results for a new type of natural human-robot interaction method based on a concept of object affordances. Finally, Sections V, VI and VII discuss and conclude the paper with suggestions for future work.

Fig. 1: Aalto University's WorkPartner robot is used as an astronaut-robot cooperation test platform.

II. HUMAN-ROBOT INTERACTION STATUS

Communication of tasks to current commercial robots resembles in many ways communication with computers. The requests to the robot are given with scripts or through graphical user interfaces. This requires the user to be both familiar with the user interfaces as well as to be able to read a display and use some type of keyboard for the input.

This kind of task communication with the robot might be still acceptable when the frequency and the number of given tasks is relatively small. This is the case for assembly robots working in factories as they perform same tasks long periods of time after the task has been given once.

When the robot is instead meant to assist us like any other person that we have become accustomed to being assisted by, the requirements of the task user interfaces change dramatically. The user might need to communicate frequently different types of requests without direct access to any kind of user manuals. The key requirement of the communication is however essentially the same. The task communication should not at least take more time and effort than if the person would do the requested task by himself.

The problem examined in this paper is to find task communication methods that would enable astronaut to communicate tasks to an assisting robot in a fast and effortless manner, both in a physical and mental respects. This includes both expected tasks that are performed according to a given mission, as well as, tasks that need to be requested due to unexpected events, for instance, due to an accident.

One important point that should be differentiated here is that the problem is not to make the assisting robots more intelligent but to improve the robots' task communication capabilities when cooperating with astronauts. Solution to the task communication problem can be in this way seen to be development of external user interface modules that are located between the robot and the human to convert the requested tasks into a format that can be directly used by the robot.

We are not also really concerned whether the robot can understand deeply the concepts that we are communicating with the task, but it should be able to understand what it is expected to do. For example, if the robot is requested to grasp a wooden object, it is probably important for the robot to only know the colour range that it needs to search for and how to grasp the object. All the other information of the properities of the wood is for example unrelevant.

III. HUMAN INSPIRED HRI METHODS

As stated before, the key issue is not to make the assiting robot itself more intelligent, but to make the task communication interfaces more intelligent. The approach in this paper is to make the user interfaces more intelligent by examining methods that are used to communicate tasks between humans and then to develop them as user interface components that enable these same methods to be used with robots. These kind of interfaces that mimic human-human communication are in this paper referred to as natural interfaces.

This section presents and analyses the usefulness of four different approaches that have been taken to develop natural human-robot interfaces. The common element in these approaches is the way the user interfaces have been inspired by task communication between humans. There are many other human-robot interaction aspects, such as social robots, that are not examined here because there is practically no empirical evidence available to argue for their usefulness [6].

For each of the presented methods a system level user interface module diagram is shown in order to clarify what exactly are the presented methods doing in practise. A general view of such module diagram is shown in Fig. 2. Essentially the user interface module is taking care of the human communication and the transforming of information to the form that can be used by the robot.

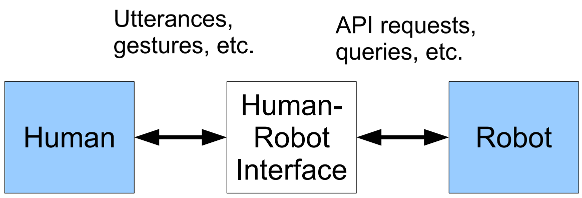

Fig. 2: A general user interface module diagram.

III.I. Peer-to-peer Dialogue

Human assistant is rarely only blindly taking and performing requests but is instead observing actively the situation and considering the correctness of the communication. If something is not as it is supposed to be, humans do not just stop but instead decide if a new dialogue should be initiated to solve the problem. This type of communication dialogue is referred to as peer-to-peer dialogue, meaning that communicating actors are being considered more or less equals as they are both able to initiate dialogues.

This type of human-human inspired dialogue has been developed to interact with robotic actors [7, 8]. The dialogue system idea is to enable robots to be human peers as they can use the humans as a resource by asking question while executing tasks, similar to the way people do. In this way the dialogue can be seen to enable the use of both human and robot capabilities when they are most appropriate. Robots are good, for example, at structured decision making and in repetitive work while humans are better in unstructured decision making, object recognition and situation assessment.

The peer-to-peer dialogue idea was first tested in an office environment with a teleoperated exploration robot [7]. The test evaluation indicated that dialogue helped especially to understand the problems that the robot tried to solve. However, at least when the human actor is in the front of a teleoperation station and focused to the robot, it did not seem to be very necessary for the human to ask questions from the robot as all the information was already available. Requesting tasks from the robot was instead, of course, a very important part of the dialogue.

The main advantage that a peer-to-peer dialogue system can provide can be said to be the sharing of information and capabilities of the actors. The robots can utilise the superior human cognitive capabilities and the humans can utilise the robot's capability to negotiate about requested tasks' parameters.

Some of the disadvantages of the dialogue have been found to be too frequently asked questions and possibly unrelevant questions [7]. This indicates that the robot's threshold for asking questions from the human actor should be probably somehow adjustable. The dialogue system performance is also very dependent on the robot's capability to evaluate if the human actor should be addressed or not.

The peer-to-peer system could be implemented also as an external user interface module, as shown in Fig. 3. In the case when the robot is requested to perform a task, the peer-to-peer user interface module can check if all the parameters required by the robot were given, and if not, they can be asked through the dialogue. In case the robot needs help instead, the peer-to-peer user interface module can initiate the dialogue to the human in case the event importance exceeds the currently used threshold level.

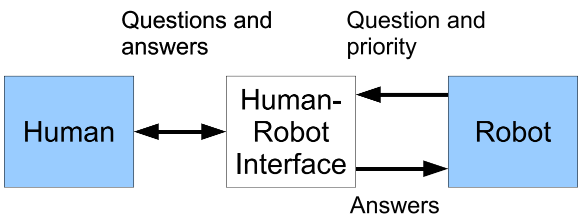

Fig. 3: Peer-to-peer dialogue implementation as an user interface module.

III.II. Perspective Taking

One distinctive feature of human-human communication is the description of spatial locations relative to other actors or objects. For example, objects can be described to be on top of other objects or on a certain side of the questioner. The exact coordinates, that are usually required by the robot, are instead in practise never used. In fact, according to the analysis of two astronauts training for an International Space Station (ISS) mission, 25% of the time the astronauts had to take the other person's perspective into consideration when performing the trained tasks [8].

Perspective-taking inspired human-robot interaction has been researched and tested in a few different applications [8, 9]. The basic idea is however the same: perspective-taking enables the robot to reason and simulate the world from the perspective of others. With perspective-taking the robot can constrain the possible action options that the user could refer to in the task communication, for instance, based on the objects visibility to the user.

The capability to understand perspectives have been implemented and tested in complex real-world tests [9]. The robot was shown, with 20 different trial runs, to be able to simulate the object visibilities from the other persons' perspectives and using different actor and object reference frames to make correct decisions about, for instance, which cone the person might be exactly referring to.

The main advantage with the system is the added flexibility to describe spatial locations. Although the descriptions can be quite rough, they are in most cases enough to constrain the options to one unambiguously defined object.

The disadvantage of the system is that the robot has to maintain a relatively correctly updated model of the environment with moving actors and objects. This can be a heavy task especially if the environment is complex. This kind of model would be however needed in some form in any case, if the target would be described from the robot's perspective, so the disadvantage is quite marginal.

A perspective taking module needs basically only information about object and actor locations in order to be able to tell the robot what exactly needs to be done. Such an environment map interface is often readily available so the perspective taking functionality could be implemented as an external user interface module. A system level view of how pespective taking could be implemented as external user interface module is shown in Fig. 4.

III.III. Common Ground

Shared knowledge and beliefs have been identified for long to be fundamental requirement of successful communication between humans [10, 11]. This assumed mutual information is referred to as common ground and the process of establishing it is referred to

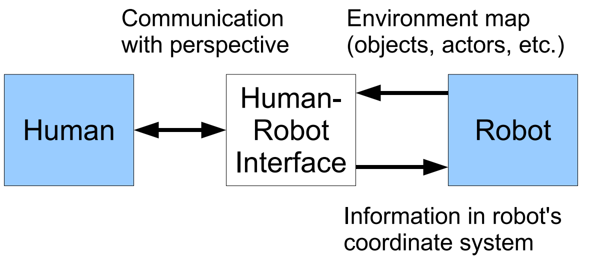

Fig. 4: Perspective taking implementation as an user interface module.

as grounding. For example, the utterance "this place is goal" can be used to create shared knowledge about a location, called goal, which can be then later used in the communication.

Approaches to establish common ground have been implemented into robots to improve the task communication [12, 13, 14]. The goal of these implementations is to have set of mutual information that can be used then to communicate about the task. With a task relevant common ground the amount of communication is minimal while still being unambiguous about the task to be performed.

[12] presents an implementation of a robot that is taught to understand basic concepts in a private house, such as kitchen and a favourite cup. These mutually understood concepts are then successfully used to give tasks to the robot. The grounding can also occur during the task communication, instead of being performed separately in advance. [14] introduced a robot that builds common ground between the user and robot by asking further specifying questions about the given task plans in cases where the mutual understanding is not completely clear. The performed tests showed that such a grounding process helped to decrease the amount of erroneous task plans given to the robot.

The main advantage of common ground is the decreased amount of task communication required. This is because the amount of common ground with a robot affects directly to how much people need to communicate with the robot [15]. If it cannot be assumed that the robot knows, for example, the names of the rooms, then the communication has to work on more general and abstract concepts, which then increases the communication effort.

One difficult issue in establishing the common ground is to know how to spend just the right amount of time to ground all the required mutual information. If a task needs to be communicated only once, then it does not make sense to use too much effort to establish the common ground. For example, people do not start to teach a tourist about places in the city when explaining directions but only use the more general descriptions. The challenge here is to find the optimal balance between effort and time used for grounding and for the task communication.

What is required of a robot module that can do grounding and utilise the common ground in task communication? It needs, in the end, only a model of current situation that the robot has available. The common ground module can then update this model, for example by naming certain locations, or read the model to transform the task communication to the format that the robot can understand. The peer-to-peer type of communication, described in Section III.I, could be used, for example, to acquire the missing information. A system level view of how usage of common ground could be enabled by implementing an external user interface module is shown in Fig. 5.

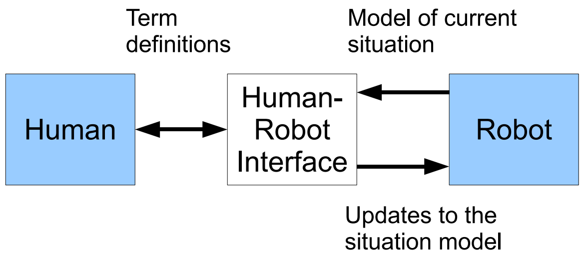

Fig. 5: Commond ground module implementation as an user interface module.

III.IV. Deictic Terms and Gestures

Terms, such as "that" or "there", which meaning depend of the current situation, are also part of human-human communication [16]. These so called deictic terms or references are not unambiguous alone but require some other completing information to be given [17]. This completing information can be, for instance, given through deictic gestures [17], such as a gaze or finger pointing, or through analysis of the situation [16, 18], such as previous utterances.

The deictic terms' and references' frequency of use in human-human communication has been found to be important especially in certain types of situations. In a situation where speech utterances and pointing gestures were allowed to explain wiring of network equipment, over 90% of the time pointing gestures were used and about 40% of these times they were also accompanied by deictic speech utterances [19]. Another test showed that over 50% of spontaneous hand movements would be deictic gestures, i.e. to be indications towards objects or actors, in a case when person is describing a painting to the other human without direct visual contact [20].

Deictic references have been for long a topic of research both in human-computer interaction [21] and in human-robot interaction [22]. The overall goal can be seen to be the completion of the ambiguous deictic task communication terms with deictic gestures. Deictic gestures have been shown to be preferred over other types of communication methods, such as speech descriptions [19].

Most of human-robot research with deictic communication has focused on communicating tasks to the robot. For example, a wheel attaching task was tested in an astronaut-robot interaction context, showing that a functional implementation using the deictic referencing can be implemented and usable [23]. The deictic gestures have been also, however, implemented to enable the robotic actor gesturing. [24] presents an experiment showing a robot with capability to gesture targets to the human actor along with verbal communication.

The main advantage of deictic gestures is that they provide mechanism to make the task communication unambiguous, most typically speech [22, 23]. Without deictic gestures the ambiguous deictic references in the communication would have to be replaced, for example, with verbal spatial descriptions. Deictic terms, such as "this", on the other hand provide a way to directly link the deictic gestures to other communication.

It is not, nevertheless, trivial to extract accurately the deictic gestures without implementing relatively complex sensor mechanisms [25]. This added system complexity means in practise vulnerability that indicates that deictic references should not be the only communication method but rather just an assisting method. The pointing gestures do not also contain any direct distance measurements so they are usually only enough to constrain the pointing to some set of possible targets.

Deictic gestures are fundamentally information about spatial relations and locations. This means that a deictic user interface module capable to use deictic gestures needs only this spatial information, in addition to the task communication containing deictic terms, as an input in order to define the task unambiguously, as shown in Fig. 6. This type of deictic references module could, however, require also the perspective taking module, described in Section III.II, because the deictic references are usually given from the point of view of the speaker [26].

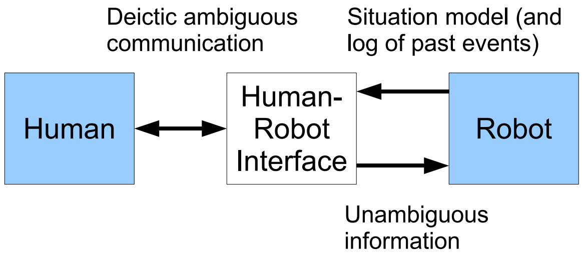

Fig. 6: Deictic references module implementation as an user interface module.

IV. OBJECT AFFORDANCES

Human-human communication as a whole is so complex that it cannot be fully understood in practise by any robot in the near future. On the other hand, humans essentially need to communicate in this versatile and flexible manner and cannot be, for example, expected to remember tens of fixed communication utterances that the robot could instead easily interpret [27].

One way to overcome these two constraints is to introduce human inspired communication methods to the robot in order to add some of the flexibility required by humans to the task communication. In here the examined approach is to mimic the human capability to associate actions to objects and use it for human-robot task communication.

IV.I. Concept of Object Affordances

The communication approach examined in this paper is based on the observation that the objects that we use are directly defining, or at least constraining, the possible actions that we can perform with those objects. This approach is mainly derived from the theory of affordances [28], which defines affordance as "action possibilities in the environment in relation to the action capabilities of an actor". In other words, the theory of affordance proposes that all objects have a property called affordance that defines which actions are possible in relation to the actors [29].

This idea of object-action relationship has been identified and further studied also in the field of psychology. It has been shown that human perception of objects enables a direct association to the possible actions that can be performed with those objects [30, 31]. Furthermore, it is worth noting that no other indication of action, than the object itself, is required for the object action-association to occur, and that the perception can be also other form than visual perception for the object-action association to work [30]. In addition, the object does not need to be visible to the user at the time of the action selection [31].

Thus, if a perception of object is received, e.g. through a pointing gesture, the human gets a number of possible actions that could be performed related to the object. This way the object alone can be used, in some cases, to communicate also the possible actions [32]. Especially when the actors capabilities are quite limited, and the context of work is possibly known, the actions associated with the object can define clearly the desired action.

IV.II. Affordance in Robotics Research

The object action relationship has been already successfully used for human action recognition, both in order to recognise objects based on actions [33] and in order to recognise actions based on objects [34]. The action recognition problem difficulty has been identified to increase proportional to the number of possible actions and objects in the environment [34]. These constraints also apply when using the object-action relationship to communicate possible actions. However, these approaches clearly demonstrate, in a concrete level, that objects can be used to infer possible actions, and vice versa.

The affordance concept has been utilised also to develop a robot subsystem that enables the robot to perceive object affordances from the environment [35]. These object affordances are determined primarily using their spatial relationships in the environment. The idea is to enable the robot to understand general functional concepts such as “chair” or “desk” from the human-robot communication when operating in a previously unknown environment. These kind of automatically obtained object affordances have been previously used, for instance, for autonomous robot control [36]. In any case, this is an example of an approach to automatically obtain the object-action associations needed also for the affordance based task communication.

IV.III. Object Affordances in Task Communication

Performing given actions to specified objects is the centre issue in Human Robot Interaction (HRI) [37]. The main goal of human-robot communication is to transfer these two coupled parameters between the human and robotic actors. The concept of object affordances has thus been long, in one form or another, at the core of robotics research.

The overall goal here in utilising object affordances is to make the communication between the astronaut and the astronaut’s assisting robot more easier. Essentially, the astronaut-robot communication should be made more intuitive, i.e. self evident to use and easy to learn, and more tolerant to the user errors.

The context in which the object affordances are researched in this paper is astronaut-robot planetary exploration. The idea is that the astronaut and robot are located in a shared work space on a planetary surfaces, such as on the Moon or Mars. This means that the human is operating in a constrained environment and is most likely able to use to some extent only speech and gestures for task communication.

Essentially the use of object affordances in task communication enables the use of implicit human-robot communication. This means that the human does not need to define explicitly both the action and target parameters of the task but the robot is able to interpret the task indirectly using the object-action associations. The direct advantage of implicit communication is that the amount of information that is required to be transfered is decreased. The disadvantage is that the robot has to have some additional information of the situation, which is in this case the list of actions that the robot can perform with different objects. A system level view of how usage of object affordances could be implemented as external user interface module is shown in Fig. 7.

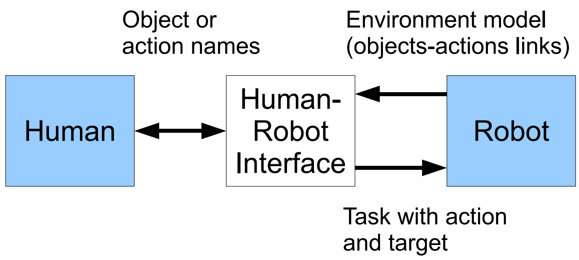

Fig. 7: Object affordances usage implementation as an external user interface module.

IV.IV. Object Affordances Test Setup

The object affordances have been tested in a geological exploration context for communicating tasks [38]. The goal of the test was to find out if the user experienced workload was decreased when using an implicit task communication method compared to an explicit task communication method. For example, “analyse rock” is an example of explicit task communication while "rock" and "analyse" are examples of implicit communication methods. In addition the test persons communication usage preferences were measured.

The idea of the test was to ask the test persons to analyse rocks in the environment and set up measurement units to the environment. The tasks could be requested either by communicating the full action and verb explicitly in the request, or by communicating the action or verb as an implicit task request. In the latter case the robot did then the object-action association in order to fully define the requested task. The test setup is described in more detail in [38].

IV.V. Object Affordances Test Results

The performed test showed that the use of object-action associations for implicit task communication was able to decrease the human workload compared to explicit task communication [38]. There was not however significant difference between the usage of action or object names as the type of implicit task communication.

However, probably the most important observation was that the test persons chose to utilise the object-action associations for task communication when they were given a free possibility to do so. As seen in Fig. 8, there was no clear preferences towards any of the three communication methods. This indicates that the object-affordances are a meaningful additional way to communicate tasks to the robot but should not be most likely the only way to do the communication.

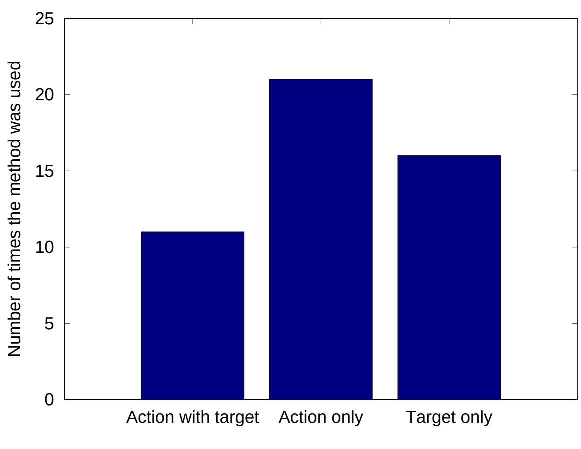

Fig. 8: Different communication methods usage in the free to choose part of the test.

V. DISCUSSIONS

The review of different human inspired communication methods showed that there is quite significant effort to add human like communication capabilities to robots. The communication methods' usefulness depends always of the examined task context but still some general guiding principles can be extracted.

Probably the most fundamental human inspired communication approach is the capability to establish common ground. Without such a capability it is not possible to extend the communication beyond a priori known objects. Adding such capability to establish common ground does not also introduce much new complexity so it is probably the most important human inspired communication method to be included in an astronaut-robot cooperation system.

Use of dialogue to make the robot more human peer like is another human inspired component that is relatively easy to add to any human-robotic system. The gained benefit from this approach is not however always self-evident because it is not always easy for the robot to know that to which extent the human can be disturbed. For certain cases, such as life-threatening situations, the possibility to initiate dialogue would most likely prove to be always advantageous.

The capability to use indirect task communication by utilising object affordances is equally a communication method that is relatively easy to include in any robotic system. This is because the only required information is essentially the list of actions that the robot can do with different objects. The explicit communication of tasks is always available in parallel so it can be always used in case implicit communication with object affordances fails. The main advantage is however additional flexibility that can, for example, help in remembering of the task communication utterances. The disadvantage is that in the presence of ambiguous object-action associations the system requires usage of some type of dialogue or situation context understanding in order to make the task communication unambiguous.

The other two examined human inspired communication methods were perspective taking and use of deictic terms in the communication. The main disadvantage of both of these methods is the need to have a relatively updated model of the current environment. Without a correctly updated model of the environment the robot reasoning will essentially produce incorrect results. This means that these methods might not provide any added value in some situations.

The perspective taking can however still be assumed to work correctly at least in some simple cases, such as "on your right side". Perspective taking is usually an additional way to communicate information. This means that in practise perspective taking failures do not finally cause really serious damage to the communication because there is always an alternative ways to communicate. The same applies also to the use of deictic terms as they can be also in the worst case left for example unused.

VI. CONCLUSIONS

Natural human-robot interaction is a way to make the robot understand communication that we humans use. The main advantage of natural human inspired interaction is that the communication is intuive, that is, it is both easy to remember and learn, because we are already familiar with it.

This paper presented five different human inspired communication methods and examined their usefulness in an astronaut-robot planetary exploration context. All of the examined communication methods can be, in principle, implemented independently as additional user interface modules. The communication methods mostly only add value to the system in the form of communication flexibility as they all are additional approaches and do not replace any existing parts.

The communication intuitiveness, implementation modularity, and added flexibility all together indicate that the natural human-robot interaction is a good approach to make the astronaut assistant robots more usable for the astronauts.

VII. FUTURE WORK

The next step in the SpacePartner project is to examine the use of object affordances in the presence of ambiguous object-action associations. The goal is to research what kind of mechanisms would be able to constrain the ambiguous object-action associations to unambiguously define the task. This kind of mechanism could be for instance peer-to-peer dialogues or modules utilising information about the current work context.

References

[1] Terrence W Fong and Illah Nourbakhsh. Interaction challenges in human-robot space exploration. *ACM Interactions*, 12(2):42–45, March 2005.

[2] Jeffrey Ferketic, Loel Goldblatt, Edward Hodgson, Sean Murray, Robert Wichowski, Arthur Bradley, Terrence W Fong, John Evans, Wendell Chun, Randy Stiles, Michael Goodrich, and Aaron Steinfeld. Toward human-robot interface standards i: Use of standardization and intelligent subsystems for advancing human-robotic competency in space exploration. In *SAE 36th International Conference on Environmental Systems*, number 2006-01-2019, July 2006.

[3] Seppo Heikkilä, Frederic Didot, and Aarne Halme. Centaur-type service robot technology assessment for astronaut assistant development. In *Proc. 10th ESA Workshop on Advanced Space Technologies for Robotics and Automation (ASTRA)*, Noordwijk, Netherlands, 2008.

[4] Aarne Halme, Ilkka Leppänen, Jussi Suomela, Sami Ylönen, and Ilkka Kettunen. WorkPartner: interactive Human-Like service robot for outdoor applications. *The International Journal of Robotics Research*, 22(7-8):627–640, July 2003.

[5] M. Heikkilä, S. Terho, M. Hirsi, A. Halme, and P. Forsman. Using signs for configuring work tasks of service robots. *Climbing and Walking Robots*, pages 909–916, 2006.

[6] T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and autonomous systems*, 42(3-4):143–166, 2003.

[7] Terrence W Fong. *Collaborative Control: A Robot-Centric Model for Vehicle Teleoperation.* PhD thesis, Robotics Institute, Carnegie Mellon University (CMU), Pittsburgh, PA, USA, 2001.

[8] T. Fong, I. Nourbakhsh, C. Kunz, L. Fluckiger, and J. Schreiner. The peer-to-peer human-robot interaction project. In *AIAA Space 2005 Conference*, Long Beach, CA, USA, 2005.

[9] J.G. Trafton, N.L. Cassimatis, M.D. Bugajska, D.P. Brock, F.E. Mintz, and A.C. Schultz. Enabling effective human-robot interaction using perspective-taking in robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 35(4):460–470, 2005.

[10] Lauri Karttunen and Stanley Peters. Conventional implicature. *Syntax and Semantics*, 11:1–56, 1979.

[11] H. H Clark and S. E Brennan. Grounding in communication. *Perspectives on socially shared cognition*, 13:127–149, 1991.

[12] S. Li. *Multi-modal interaction management for a robot companion.* PhD thesis, Bielefeld University, 2007.

[13] K. Stubbs, D. Wettergreen, and P.J. Hinds. Autonomy and common ground in human-robot interaction: A field study. *IEEE Intelligent Systems*, 22:42–50, 2007.

[14] K. Stubbs, D. Wettergreen, and I. Nourbakhsh. Using a robot proxy to create common ground in exploration tasks. In *Proceedings of the 3rd*

ACM/IEEE international conference on Human robot interaction, pages 375–382. ACM, 2008.

[15] S. Kiesler. Fostering common ground in human-robot interaction. In *IEEE International Workshop on Robot and Human Interactive Communication (ROMAN)*, pages 729–734, 2005.

[16] W.J.M. Levelt, G. Richardson, and W. La Heij. Pointing and voicing in deictic expressions. *Journal of Memory and Language*, 24(2):133–164, 1985.

[17] D. McNeill. *Hand and mind: What gestures reveal about thought*. University of Chicago Press, 1996.

[18] C. Knipping. A method for revealing structures of argumentations in classroom proving processes. *The International Journal on Mathematics Education (ZDM)*, 40(3):427–441, 2008.

[19] M. Bauer, G. Kortuem, and Z. Segall. Where are you pointing at? a study of remote collaboration in a wearable videoconference system. In *The Third International Symposium on Wearable Computers*, pages 151–158, 1999.

[20] M. Gullberg. Gestures in spatial descriptions. *Lund Working Papers in Linguistics*, 47:87–97, 1999.

[21] R.A. Bolt. Put-that-there: Voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, pages 262–270. ACM, 1980.

[22] P.K. Pook and D.H. Ballard. Deictic human/robot interaction. *Robotics and Autonomous Systems*, 18(1-2):259–269, 1996.

[23] A.G. Brooks and C. Breazeal. Working with robots and objects: Revisiting deictic reference for achieving spatial common ground. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, page 304. ACM, 2006.

[24] Osamu Sugiyama, Takayuki Kanda, Michita Imai, Hiroshi Ishiguro, Norihiro Hagita, and Yuichiro Anzai. Humanlike conversation with gestures and verbal cues based on a three-layer attention-drawing model. *Connection Science*, 18(4):379–402, 2006.

[25] D. Perzanowski, A.C. Schultz, and W. Adams. Integrating natural language and gesture in a robotics domain. In *Proceedings of the IEEE International Symposium on Intelligent Control: ISIC/CIRA/ISAS Joint Conference*, pages 247–252. Citeseer, 1998.

[26] J.F. Duchan. *Deixis in Narrative: A Cognitive Science Perspective*, chapter Chapter 9: Preschool Children's Introduction Of Characters into Their Oral Stories: Evidence for Deictic Organization of First Narratives, pages 227–241. Lawrence Erlbaum Associates, 1995.

[27] G. W Furnas, T. K Landauer, L. M Gomez, and S. T Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11):971, 1987.

[28] J.J. Gibson. The theory of affordances. *Perceiving, acting and knowing*, pages 67–82, 1977.

[29] J. McGrenere and W. Ho. Affordances: Clarifying and evolving a concept. In *Graphics Interface*, pages 179–186, 2000.

[30] J. Grezes and J. Decety. Does visual perception of object afford action? evidence from a neuroimaging study. *Neuropsychologia*, 40(2):212–222, 2002.

[31] M. Tucker and R. Ellis. Action priming by briefly presented objects. *Acta Psychologica*, 116(2):185–203, 2004.

[32] K. Tylén, M. Wallentin, and A. Roepstorff. Say it with flowers! An fMRI study of object mediated communication. *Brain and language*, 108(3):159–166, 2009.

[33] P. Peursum. *Using Human Activity to Indirectly Recognise Objects in Indoor Wide-Angle Scenes*. PhD thesis, Curtin University of Technology, 2005.

[34] D.J. Moore, I.A. Essa, and M.H. Hayes. Exploiting human actions and object context for recognition tasks. In *IEEE International Conference on Computer Vision*, volume 1, pages 80–86, 1999.

[35] R. Moratz and T. Tenbrink. Affordance-based human-robot interaction. *Lecture Notes in Computer Science*, 4760:63–76, 2008.

[36] E. Sahin, M. Cakmak, M.R. Dogar, E. Ugur, and G. Ucoluk. To afford or not to afford: A new formalization of affordances toward affordance-based robot control. *Adaptive Behavior*, 15(4):447, 2007.

[37] V. Klingspor, J. Demiris, and M. Kaiser. Human-robot communication and machine learning. *Applied Artificial Intelligence*, 11(7):719–746, 1997.

[38] S. Heikkilä, A. Halme, and A. Schiele. Human-human inspired task and object definition for astronaut-robot cooperation. In *Proc. of The 10th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS)*, 2010.