Common Ground in Human-Robot Interaction

Tom Gross
Human-Computer Interaction Group
University of Bamberg
96045 Bamberg, Germany
<firstname>.<lastname>(at)uni-bamberg.de

***Abstract*—The efficiency of social interaction depends on common ground—that is, a mutual understanding of the actors' behavior and their environment. Insufficient common ground can have severe negative effects such as misunderstandings, conflicts, or event damage. In this position paper I introduce my perspective common ground as both a basis and a result of a mutual understanding of each other's actions and intentions; and I point out and sketch how awareness research from Computer-Supported Cooperative Work might contribute.**

Keywords—Recognition of Actions and Intentions; Computer-Supported Cooperative Work; Common Ground; Awareness.

I. Introduction and Motivation

In social interaction it is important to adequately understand each other's actions and intentions. Humans have a natural ability to infer intentions from actions [2]. Besides capturing and analysing the actions per se, humans also scan and infer from facial expressions and gaze, etc. [17]. Likewise, successful human-robot interaction requires a mutual understanding of actions and intentions between the human actor or actors and the robot or robots. Furthermore, both humans and robots need to understand their shared environment [3]. One of the big challenges here is the question whether robots can have intentionality as a point of departure for intentions and consequently for actions. In the existing literature there have been great discussion and analyses towards a better understanding of the roles that the body and the intentionality of the actors play [18]. This is very central since it is assumed that from an intentional stance only humans have intentionality [17].

II. Common Ground

Independent of similarities and differences between humans' and robots' intentionality, in successful human-robot interaction some minimal common ground is required [3] [10]. Common ground can have different categories that have been identified for human communication and beyond [4, 13]: conversational conventions: these are the very basic agreements of the style of the social interaction (e.g., turn-taking); communal common ground: these are basic assumptions the actors have amongst each other (e.g., that the language of a dialogue is English); personal common ground achieved before the conversation: these are a-priori agreements (e.g., the human actor knows and the robot is programmed to perform their shared task in car production); as well as personal common ground developed during the conversation: these are outcomes of the mutual interaction (e.g., the robot might be able to learn new words or pronunciations of the human actor).

A very central lesson here is that the common ground is a mutual construct—that is, all human actors and robots involved require this understanding of each other. And furthermore it is a recursive construct—that is, all parties involved require an understanding of the others' understanding [13].

Having this reciprocity in mind we can furthermore distinguish the degree of autonomy of the parties involved (see also [16]). Both the human actor as well as the robot can either be dependent on the other party (e.g., the user requiring the robot to carry and transport a heavy object), or independent of it (e.g., the human and the robot acting in the same space, but not directly interacting with each other). Figure 1 illustrates the user's and robot's dependencies and the corresponding types of shared understanding (the two arrows illustrated that the dependencies can also change over time).

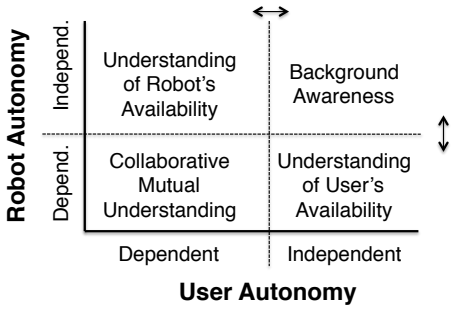

Fig. 1. User's and robot's dependencies and corresponding types of shared understanding (*Source*: inspired by [16]).

III. Towards Designing for Common Ground

In order to support common ground it is important that the design of the human-robot interaction addresses these types of dependencies and respects this mutual and recursive nature of common ground.

In the field of Computer-Supported Cooperative Work (CSCW) there has been interesting research on technical support for common ground in human-computer-human interaction. Traditionally, the concept here has been referred to as *awareness*, which can be defined as 'a user's internal knowing and understanding of a situation including other users and the environment that is gained through subtle practices of capturing and interpreting information' [7, p. 432].

Overall CSCW offers considerable insight with respect to awareness research. Starting with early ethnographic studies of mutual heeding of team members in situations of tight collaboration and early systems to capture and present

awareness information it eventually evolved into an understanding and support for specific types of awareness such as coexistence awareness (on the mutual presence of users); and on cooperative awareness in shared workspaces and in group editors [7].

Recently research in CSCW has pointed out the importance of *shared intentionality* in human-computer-human interaction [15]. Tenenberg et al. point out that: 'a number of philosophers have defined notions of intentionality that extend beyond the individual … "we-intentionality" … also called shared cooperative activity, shared intentionality, joint intentionality, and collective intentionality' [15, p. 240]. Very often in the classical literature on phenomenology that dealt with intentionality the focus has been on individuals—especially in the literature of the founder of phenomenology Husserl [11], but also in his students and followers such as Heidegger [9] and Merleau-Pointy [12]. It was mostly Schuetz who started including a strong social perspective [14].

One factor that increases the complexity of mutually understanding the actions and intentions is the finding that humans in any kind of social interaction have a tendency to actively present themselves. For instance, Hancock et al. could show that users in instant messaging who get mutual information of each other's online presence sometimes use deception to either avoid or to get out of on-going conversations [8]. They use the term *butler lies* in comparison to the behaviour of butlers who sometimes are instructed to send visitors away despite the presence of the host. Goffman studied social interaction for several decades and developed a framework of social interaction [1, 6]. In this framework he uses the metaphor of a theatre performance pointing out that humans in social interaction always try to make an impression on the other participants. This has not necessarily anything to do with cheating, but very often is related to presenting oneself differently in specific social contexts (e.g., at work versus in the sports club). Still, when trying to understand the behaviour and intentions behind activities it is essential to keep in mind that the motivation of the actor can be two-fold: the intentions towards the task at hand, and the desire to influence one's appearance.

IV. Conclusions and Summary

In his book entitled 'Where the Action Is: The Foundations of Embodied Interaction' Dourish points out that embodied interaction has a bodily tangible dimension and a social dimension. He writes: 'embodied interaction is interaction with computer systems that occupy our world, a world of physical and social reality, and that exploit this fact in how they interact with us.' [5, p. 3]. Environments populated with humans and robots are increasingly based on such embodied interaction.

In this workshop on 'The Role of Intentions in Human-Robot Interaction' at HRI 2017 I would like to contribute insight from research in CSCW on awareness and discuss overlaps, contradictions, and especially synergies between different fields interested in a better understanding and technical support for the recognition of action and intention in human-human and human-robot scenarios.

References

[1] Beckmann, C. and Gross, T. Social Computing - Bridging the Gap Between the Social and the Technical. In Proceedings of the 6th International Conference on Social Computing and Social Media - SCSM 2014 (June 22-27, Crete, Greece). Springer-Verlag, Heidelberg, 2014. pp. 25-36.

[2] Blakemore, S.-J. and Decety, J. From the Perception of Action to the Understanding of Intention. Nature Reviews, Neuroscience 2 (Aug. 2001). pp. 561-567.

[3] Chai, J.Y., She, L., Fang, R., Ottarson, S., Littley, C., Liu, C. and Hanson, K. Collaborative Effort Towards Common Ground in Situated Human-Robot Dialogue. In Proceedings of the 9th ACM/IEEE International Conference on Human-Robot Interaction - HRI 2014 (Mar. 3-6, Bielefeld, Germany). ACM, N.Y., 2014. pp. 33-40.

[4] Clark, H.H. and Brennan, S.E. Grounding in Communication. In Resnick, L.B., Levine, J.M. and Teasley, S.D., eds. Perspectives on Socially Shared Cognition. American Psychological Association, Washington, DC, 1991. pp. 127-149.

[5] Dourish, P. Where the Action Is: The Foundations of Embodied Interaction. MIT Press, Cambridge, MA, 2001.

[6] Goffman, E. The Presentation of Self in Everyday Life. Doubleday Anchor Books, N.Y., 1959.

[7] Gross, T. Supporting Effortless Coordination: 25 Years of Awareness Research. Computer Supported Cooperative Work: The Journal of Collaborative Computing 22, 4-6 (Aug.-Dec. 2013). pp. 425-474.

[8] Hancock, J.T., Birnholtz, J.P., Bazarova, N., Guillory, J., Perlin, J. and Amos, B. Butler Lies: Awareness, Deception, and Design. In Proceedings of the Conference on Human Factors in Computing Systems - CHI 2009 (Apr. 4-9, Boston, MA). ACM, N.Y., 2009. pp. 517-526.

[9] Heidegger, M. Sein und Zeit. Elfte, unversaenderte Auflage. Max Niemeyer Verlag Tuebingen, Tuebingen, 1967.

[10] Herr, S., Gross, T., Gradmann, M. and Henrich, D. Exploring Interaction Modalities and Task Allocation for Household Robotic Arms. In Conference on Human Factors in Computing Systems - CHI 2016 (May 7-12, San Jose, CA). ACM, N.Y., 2016. pp. 2844-2850.

[11] Husserl, E. The Basic Problems of Phenomenology - From the Lectures, Winter Semester, 1910-1911 (Grundprobleme der Phaenomenologie, Aus den Vorlesungen, Wintersemester 1910/1911 in 'Zur Phaenomenologie der Intersubjektivitaet, Husserliana XIII, Iso Kern (ed.), Trans. Ingo Farin and James G. Hart). Springer-Verlag, Heidelberg, 2006.

[12] Merleau-Ponty, M. The Phenomenology of Perception (Phenomenologie de la Perception; Original Work Published in French in 1945, Trans. Colin Smith). Routledge (Gallimard), N.Y. (Paris), 1962 (1945).

[13] Monk, A. Common Ground in Electronically Mediated Conversation. Morgan & Claypool Publishers, San Rafael, CA, 2009.

[14] Schuetz, A. The Phenomenology of the Social World (Der sinnhafte Aufbau der sozialen Welt: eine Einleitung in die verstehende Soziologie; Original Work Published in German in 1932, Trans. George Walsh and Frederick Lehnert). Northwestern University Press, Evanston, 1967 (1932).

[15] Tenenberg, J., Roth, W.-M. and David, S. From I-Awareness to We-Awareness in CSCW. Computer Supported Cooperative Work: The Journal of Collaborative Computing (JCSCW) 25, 4 (Oct. 2016). pp. 235-278.

[16] Vernon, D., Lowe, R., Thill, S. and Ziemke, T. Embodied Cognition and Circular Causality: On the Role of Constitutive Autonomy in the Reciprocal Coupling of Perception and Action. Frontiers in Psychology 6, 1660 (Oct. 2015). pp. 1-9.

[17] Wykowska, A., Kajopoulos, J., Obando-Leiton, M., Chauhan, S.S., Cabibihan, J.-J. and Cheng, G. Humans are Well Tuned to Detecting Agents Among Non-Agents: Examining the Sensitivity of Human Perception to Behavioural Characteristics of Intentional Systems. International Journal of Social Robotics 7 (May 2015). pp. 767-781.

[18] Ziemke, T. The Body of Knowledge: On the Role of the Living Body in Grounding Embodied Cognition. BioSystems 148 (Oct. 2016). pp. 4-11.