Movement coordination in repetitive joint action tasks: Considerations on human-human and human-robot interaction*

Tamara Lorenz *Student Member, IEEE*, and Sandra Hirche *Member, IEEE*

***Abstract*— It is a common idea that robotic design for human-robot interaction can benefit from approaches taken from human joint action research. In this position paper a pre-defined example is operationalized to shed light on recent findings in human movement coordination and turn-taking in repetitive tasks. Both topics are also considered for human-robot interaction. The paper closes with a discussion on open questions and unsolved problems that are not considered in human psychological research but play a major role when a transfer to robotics is intended.**

I. Introduction

Human-Robot Interaction (HRI) as a field of research is getting more and more attention among researchers in both robotics and psychological research. The difficulty here is twofold: on the one hand, engineers and designers of robots often start by realizing a technically possible design but do not consider the actual needs and expectations of a human that has to interact with the robot. On the other hand, researchers in psychology tackle the joint action or interaction approach from a very fundamental point of view – which sometimes lacks the direct applicability into the robot design process. To provide some insight in actual problems that arise from this discrepancy a special case scenario was chosen: a simple joint pick and place task. In Section II and III this scenario and necessary assumptions for the following argumentation are outlined. In this position paper the focus is on behavioral differences that are established in reaction towards the behavior of the respective other actor. By tackling the behavioral differences between a human interacting with another human and a human interacting with a robot, it sheds light on the requirements a human-robot joint action task has to meet.

Finally we will argue that movement synchronization (MS) as a simple but basic phenomenon of human interaction might be a good way to improve HRI.

II. The Example Scenario

In the example which was proposed by the workshop organizers [1], a human and a robot, are standing at a table. Each actor has two numbered bricks and one triangle, see Figure 1a. Their task is to jointly build a pile from the bricks in front of them. The bricks have to be used in numerical

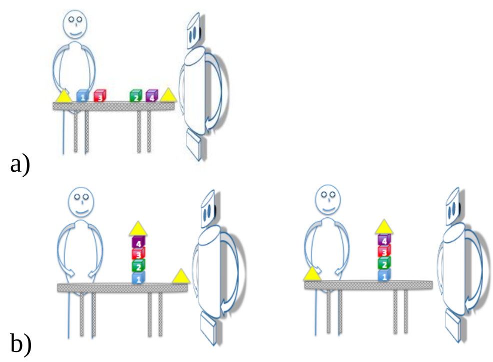

Figure 1 Example Setup with two actors, a human and a robot (taken from [1]): a) initial state; b) two potential final states; in the present interpretation the final state depends upon which actor starts as the joint action is expected to be alternately.

order. In the end, one triangle has to be put on top. From this, two possible end-states arise, see Figure 1b.

If the already well-established definition for joint action, which Natalie Sebanz gave in 2006 is taken into account, namely that " […] joint action can be regarded as any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment." [2], the example can be considered as joint action between the human and the robot.

As the intention of this position paper is to highlight the differences in performance of human-human compared to human-robot interaction, the following sections will consider both scenarios. Therefore we also consider the same task being performed by two humans. Humans and robot will also be referred to as *actors* in the following.

III. Problem Definition

Although a pick-and-place task is usually considered being a fairly simple task, it still requires a lot of different processes if the joint action should have a successful outcome. However, while these processes are cognitively also not trivial for interacting humans, they have the ability to do so while for today's robots this is still a challenge [3].

First of all, both actors have to form a common representation of the task. Then, the task has to be planned and the role of each partner has to be assigned. Still being not ready to execute, the intention of the interaction partner has to be estimated in order to plan and adjust the own part appropriately. The latter is an ongoing process which is also required for online adjustment during task performance.

Although there are also a lot of open questions regarding the action planning process both in human and in robotic

*This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) as part of the Excellence Cluster "Cognition for Technical Systems" (CoTeSys).

T. Lorenz and S. Hirche are with the Institute for Information-Oriented Control, Technische Universität München (TUM). TL is also with the Experimental Psychology Department, Ludwig-Maximilians Universität (LMU), Munich, Germany (phone: +49-89-289-25736; fax: +49-89-289-25724; e-mail:{t.lorenz,hirche}@tum.de)

research, this article focuses on the actual performance of the task. This process starts when the action planning and role assignment are already finished and both interaction partners know what to do. In order to constrain the scenario to this problem, we made some assumptions towards the given example:

(1) Each actor is responsible for his part of the task and able to fulfill his part independently, i.e. no stabilization of the pile by one actor or similar is required while the other actor puts his brick down on top.
(2) The actors already share the goal and know about the general task requirements.
(3) Both actors know who is going to set the first brick: this actor will also set the triangle on top of the pile, see Figure 1b left. Together with assumption (2) this means that the intentions of the agents are clear.
(4) In an HRI scenario, the robot is safe as the system runs robustly and reliable; colliding with the robot would not cause more harm to the human than colliding with another human.

IV. I AM ACTING, YOU ARE ACTING: SELF- AND OTHER ACTION REPRESENTATION

Considering assumption (2), the actors know about the task, the intended outcome and the necessary steps to be fulfilled by each actor. However, for being able to engage in a joint action task, each actor also has to form a representation of both one's own and the other's behavior during action execution.

A. … human-human

According to the mirror neuron theory [4], humans use their own body schema to understand the behavior of their counterparts [5]. This means, that because they expect their interaction partner to have similar abilities compared to their own, they somehow take perspective of the other person, project their own abilities into this scenario and use this information to create a representation of what the counterpart is likely to do. Framed in the present example this means, if a human and another human are working together to build the pile, they have the "tools" to form the task representation of the other person and the success will only depend on each actor's physical ability.

B. … human-robot

But what happens when the body schema doesn't match? Müller et al. [6] showed that already forming a representation of an *out-group* member might be biased (i.e. somebody who has a different skin color). However they also showed that this might be a matter of habit and conducted a second study in which they provided some prior information about the out-group member. Here the difficulties were overcome.

If now a human is confronted with a robot like in the given example, the formation of a representation might depend upon the level of physical human-like appearance of the robot [7] or to the level of training [8] a person has in the interaction with the robot (i.e. how much does the human know about the movement abilities of the robot). Thus, if there is no previous training available, and the robot doesn't have a human-like appearance, forming a representation might be difficult for the human and bias the behavior that the robot might expect due to the data he has learned from observing human (e.g. by means of motion tracking).

When building a pile of bricks in an alternating way this might in the best case lead to hesitative behavior - because forming a plan costs more effort, or to more cautious behavior and bigger deviations - because the human wants to avoid collisions, an attempt that can already be observed in human joint action compared to human single action [9]. In the worst case the human might not even dare to share a workspace with the robot [10].

C. …robot-human

The problem for the robot starts already by forming a representation of the task [11]. In order for being able to plan and re-plan its own actions, the robot needs to know geometrically where it is located in space, where the objects and the interaction partner are and what the outcome of the task should be. For the interaction itself this information has to be integrated in a higher level action plan with all information being available online. Thus, for the current example, the robot needs a system that allows for detection of the bricks' and the pile's location in relation to its own position. Furthermore, besides having a plan for its own action, it needs to be able to online detect how the environment is changing, i.e. where the human is located and how the general task is proceeding over space and time.

V. WE ARE ACTING: JOINT MOVEMENT COORDINATION AND TURN-TAKING

If both actors want to achieve the goal of a built-up pile they have to jointly work together because every actor can only fulfill part of the task. Due to the task constraints given, the actors have to alternately place a brick in the shared work space. For being able to coordinate one's movements successfully with each other, both actors constantly need to adapt and keep track of what the other is doing in order to be able to online react to unexpected behavior. Furthermore, as the task is to be fulfilled alternately, the task requires turn-taking.

A. …human-human

A very typical coordination behavior that emerges inevitably during repetitive pick-and-place tasks is some kind of rhythm formation or movement synchronization [12], [13]. Although for a long time being studied in undirected tasks like rocking in chairs [14], it was shown that this also holds true for goal-directed movements that require precise movements and bear similarities to pick-and-place tasks [15]–[17]. If an obstacle is introduced in one person's workspace, and the trajectory is slightly modified, this also causes modifications especially in the temporal behavior of affects the interaction partner [18]. Nevertheless people do jointly engage in this task and make sure to adapt to the interaction partners restrictions. Thus, even if the task is therefore more difficult, people still synchronize their movements.

Synchronization is usually given when two actors perform the same action at the same time (in-phase relation), but also when they perform complementary movements at the same time (anti-phase). Framed to the given example, this means that during in-phase MS, actors would try to put their brick at the same time. Logically that is not very promising – therefore anti-phase MS seems the way to go, i.e. the first actor moves forwards to put his brick while the second actor moves backwards to catch the next one. Anti-phase behavior of this kind is what is also termed *turn-taking*. Here the actors have to determine whose turn it is to put the brick onto the pile in the joint workspace, while the other actor can use this time to prepare for the next step.

In general humans seem to have a coordination and timer model that helps them to estimate intervals between events. [19]. Furthermore, there seems to be a coupling of intra- and interpersonal behavior during interaction to the extent that adjustments in intrapersonal coordination affect the coordination with another person, and vice versa [20]. Finally, it was shown that reducing the variability in one's own movements can be a strategy to achieve predictability for the interaction partner [21,30].

In the given example this means that people will adapt their own behavior to the behavior of the interaction partner, both in a temporal and in a spatial way. In milliseconds time they will find a rhythm for their interaction and MS is established. However if one person is setting the brick while the other person is moving backwards some deviating behavior will be performed. Both the temporal and the spatial adaptation are probably expected and incorporated into the action plan.

B. *…robot-human*

Overall, MS seems to be a promising approach for human robot interaction. Different models are already available with the goal to include synchrony and turn-taking into the repertoire of robotic behavior. Revel and Andry [22] present a neural network architecture based on coupled oscillators that by changing the coupling direction is able to produce both synchronizing and turn-taking behavior between two robots. Another approach by Kose-Bagci et al. for designing a turn-taking robot (in a drumming game) is to use probabilistic models [23]. However both of these approaches were not designed for goal-directed behavior which would be necessary for building a pile.

Using data from a human- human interaction task [16], Mörtl et al. [24] showed that goal-directed movements can successfully be replicated by attractor dynamics of coupled phase oscillators inspired by the HKB model [25]. Here, participants established in-phase as well as anti-phase relations. This model predicts the dynamics of inter-human movement coordination and can directly be implemented to human–robot interaction. Going one step further, Mörtl et al. [17] describe the goal-directed interaction task by closed movement trajectories and interpreted those as limit cycles for which instantaneous phase variables are derived based on oscillator theory. Additionally, events are introduced as anchoring points which segment these trajectories into multiple parts. Utilizing both continuous phases and discrete events in a unifying view, a continuous dynamical process is designed which synchronizes the derived modes. With such a model it is possible for a robot to synchronize its motion to the behavior of a human while taking the picking and placing of the bricks as anchoring points in space and time.

C. *… human-robot*

The developed concept in [17] was furthermore implemented to an anthropomorphic robot and successfully tested in a HRI task in which the robot performs a prototypical pick-and-place task jointly with human partners.

However, a very striking problem when it comes to MS and turn-taking in HRI is the problem of adaptation. In a study in which a human was interacting with a non-adaptive robot in a goal-directed tapping task, it was shown that humans do not take over the full adaptation load to synchronize their movements with the robot [26]. Thus, there must be something in the interaction behavior of humans that triggers the share of the joint action. This means, humans only start their own adaptation to the joint action process if they perceive the adaptation also from the other side. The tricky thing here is that even if the robot is adaptive, the amount of adaptation is still to be determined. The models developed in [17] and [24] allow for an adjustment of the amount of adaptation and might therefore be a good start for exploring the actual acceptable amount that successfully trigger the emergence of adaptive behavior also in HRI. Preliminary results show that this difference is very fine-grained and might furthermore depend on personal preferences and the level of experience with regard to the interaction with robots. Furthermore, if the adaptation from the robot is too high, this might also cause a feeling of uncannyness with regard to the robots behavior.

Overall the model from [17] seems to be a good starting point also for describing an ongoing movement interaction as in the example. It is capable of both describing the timing of the turn-taking by emergent dynamics and also account for the events of picking and placing objects like bricks on a pile. A nice side effect of using MS dynamics for robotic action is, that it does not only support turn-taking behavior but is known to also increases affiliation between interaction partners – something that should not be neglected especially when thinking about social HRI [27], [28].

VI. Discussion and Open Issues

When considering the requirements for a successful joint action between a human and a robot, it appears that these requirements are of course similar, as the task for the robot has to be planned in the biological world. Thus, it seems beneficial to understand the underlying principles of human interaction in order to trying to reproduce them with a robotic system. However much human behavior is not understood in humans either and replicating human behavior in robotics, would require to first understanding the human behavioral principles in detail. Besides, when working on the problem of jointly picking and placing objects on the intersection of human and robotics some very practical problems arise:

(1) Although we considered the robot to be safe – and even if we assume that the robot is considered to be safe by

the interacting human – one cannot assume that the human will react in exactly the same way as it would with a human interaction partner. This might depend on experience in the interaction with robots, but could maybe also be reflected in the movements of the robot.

(2) If humans interact with each other they expect a certain range of velocity and reactivity connected to the given task. If this is not reflected in the robot's behavior, it can lead to a feeling of uncertainty which might potentially lead to a reformation of the representation of the robot's abilities because its behavior was not predictable. The same holds true for the movement profile: usually the minimum-jerk movement profile is considered to be human-like and therefore provides the best standard in HRI research. However little is known about how humans actually perceive and accept this behavior with respect to its application on different robotic arms.

In both cases, MS could increase predictability by means of rhythm generation and by meeting temporal and spatial expectations. As MS is a simple but very influential phenomenon in human interaction, we think that it is a promising approach to equip a robotic system with the ability to recognize and make use of synchronous mechanisms to be able to interact with humans in a predictable and even social way. In this context, also the breakdown of MS could be interesting to study, as this is also a means for communication [29, 30]. However, as mentioned before it is very important to consider the extent to which the robot should adapt to the behavior of the human. This might on the one hand trigger reciprocity and mutual engagement but also prevent a perceived uncannyness in the robots behavior and thus increase its acceptance as partner for the joint action task.

References

[1] A. Clodic, R. Alami, E. Pacherie, and C. Versper, "Towards a framework for joint action." [Online]. Available: http://fja2014.sciencesconf.org/resource/page/id/3.

[2] N. Sebanz, H. Bekkering, and G. Knoblich, "Joint action: bodies and minds moving together.," *Trends Cogn. Sci.*, vol. 10, no. 2, pp. 70–6, Feb. 2006.

[3] J. a. Adams, "Human-Robot Interaction Design: Understanding User Needs and Requirements," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 49, no. 3, pp. 447–451, Sep. 2005.

[4] G. Rizzolatti, "The mirror neuron system and its function in humans.," *Anat. Embryol. (Berl).*, vol. 210, no. 5–6, pp. 419–21, Dec. 2005.

[5] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. C. Mazziotta, and G. Rizzolatti, "Grasping the intentions of others with one's own mirror neuron system.," *PLoS Biol.*, vol. 3, no. 3, p. e79, Mar. 2005.

[6] B. C. N. Müller, S. Kühn, R. B. van Baaren, R. Dotsch, M. Brass, and A. Dijksterhuis, "Perspective taking eliminates differences in co-representation of out-group members' actions.," *Exp. Brain Res.*, vol. 211, no. 3–4, pp. 423–8, Jun. 2011.

[7] T. Chaminade, "A Social Cognitive Neuroscience Stance on Human-Robot Interactions," in *The International Conference SKILLS 2011*, 2011, vol. 1, p. 4.

[8] M. a. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," *Found. Trends® Human-Computer Interact.*, vol. 1, no. 3, pp. 203–275, 2007.

[9] C. Vesper, A. Soutschek, and A. Schubö, "Motion coordination affects movement parameters in a joint pick-and-place task.," *Q. J. Exp. Psychol.*, vol. 62, no. 12, pp. 2418–32, Dec. 2009.

[10] A. Lenz, S. Lallee, S. Skachek, A. G. Pipe, C. Melhuish, and P. F. Dominey, "When shared plans go wrong: From atomic- to composite actions and back," *2012 IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, pp. 4321–4326, Oct. 2012.

[11] R. Petrick, D. Kraft, and K. Mourao, "Representation and integration: Combining robot control, high-level planning, and action learning," in *Proceedings of the 6th international cognitive robotics workshop (CogRob)*, 2008, pp. 32–41.

[12] J. Issartel, L. Marin, and M. Cadopi, "Unintended interpersonal co-ordination: 'can we march to the beat of our own drum?'.," *Neurosci. Lett.*, vol. 411, no. 3, pp. 174–9, Jan. 2007.

[13] R. C. Schmidt and B. O'Brien, "Evaluating the Dynamics of Unintended Interpersonal Coordination," *Ecol. Psychol.*, vol. 9, no. 3, pp. 189–206, Sep. 1997.

[14] M. J. Richardson, K. L. Marsh, R. W. Isenhower, J. R. L. Goodman, and R. C. Schmidt, "Rocking together: dynamics of intentional and unintentional interpersonal coordination.," *Hum. Mov. Sci.*, vol. 26, no. 6, pp. 867–91, Dec. 2007.

[15] T. Lorenz, B. N. S. Vlaskamp, A.-M. Kasparbauer, A. Mörtl, and S. Hirche, "Dyadic movement synchronization while performing incongruent trajectories requires mutual adaptation," *Front. Hum. Neurosci.*, vol. submitted, 2014.

[16] T. Lorenz, A. Mörtl, B. Vlaskamp, A. Schubö, and S. Hirche, "Synchronization in a goal-directed task: Human movement coordination with each other and robotic partners," in *20th IEEE International Symposium on Robot and Human Interactive Communication*, 2011, pp. 198–203.

[17] A. Mörtl, T. Lorenz, and S. Hirche, "Rhythm Patterns Interaction - Synchronization Behavior for Human-Robot Joint Action," *PLoS One*, no. accepted, 2014.

[18] T. Lorenz, B. N. S. Vlaskamp, A.-M. Kasparbauer, A. Mörtl, and S. Hirche, "Dyadic movement synchronization during unilteral obstacle avoidance requires mutual adaptation," *Front. Hum. Neurosci.*

[19] R. B. Ivry and T. C. Richardson, "Temporal control and coordination: the multiple timer model.," *Brain Cogn.*, vol. 48, no. 1, pp. 117–32, Feb. 2002.

[20] V. C. Ramenzoni, T. J. Davis, M. a Riley, K. Shockley, and A. a Baker, "Joint action in a cooperative precision task: nested processes of intrapersonal and interpersonal coordination.," *Exp. Brain Res.*, vol. 211, no. 3–4, pp. 447–57, Jun. 2011.

[21] C. Vesper, R. P. R. D. van der Wel, G. Knoblich, and N. Sebanz, "Making oneself predictable: reduced temporal variability facilitates joint action coordination.," *Exp. Brain Res.*, vol. 211, no. 3–4, pp. 517–30, Jun. 2011.

[22] A. Revel and P. Andry, "Emergence of structured interactions: from a theoretical model to pragmatic robotics.," *Neural Netw.*, vol. 22, no. 2, pp. 116–25, Mar. 2009.

[23] H. Kose-Bagci, K. Dautenhahn, and C. L. Nehaniv, "Emergent dynamics of turn-taking interaction in drumming games with a humanoid robot," *RO-MAN 2008 - 17th IEEE Int. Symp. Robot Hum. Interact. Commun.*, pp. 346–353, Aug. 2008.

[24] A. Mörtl, T. Lorenz, B. N. S. Vlaskamp, A. Gusrialdi, A. Schubö, and S. Hirche, "Modeling inter-human movement coordination: synchronization governs joint task dynamics.," *Biol. Cybern.*, vol. 106, no. 4–5, pp. 241–59, Jul. 2012.

[25] H. Haken, J. A. S. Kelso, and H. Bunz, "A theoretical model of phase transitions in human hand movements," *Biol. Cybern.*, vol. 51, no. 5, pp. 347–356, Feb. 1985.

[26] T. Lorenz, A. Mortl, and S. Hirche, "Movement synchronization fails during non-adaptive human-robot interaction," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2013, pp. 189–190.

[27] M. J. Hove and J. L. Risen, "It's All in the Timing: Interpersonal Synchrony Increases Affiliation," *Soc. Cogn.*, vol. 27, no. 6, pp. 949–960, Dec. 2009.

[28] T. Lorenz, "Synchrony and reciprocity for social companion robots: benefits and challenges," in *Taking Care of Each Other: Synchronization and Reciprocity for Social Companion Robots; Workshop at the International Conference on Social Robotics*, 2013.

[29] T. Wheatley, O. Kang, C. Parkinson, and C. E. Looser, "From Mind Perception to Mental Connection: Synchrony as a Mechanism for Social Understanding," *Soc. Personal. Psychol. Compass*, vol. 6, no. 8, pp. 589–606, Aug. 2012.

[30] G. Pezzulo, F. Donnarumma, and H. Dindo, "Human sensorimotor communication: a theory of signaling in online social interactions.," *PLoS One*, vol. 8, no. 11, p. e79876, Jan. 2013.