Time Matters - Even (more so) in Human-Robot Interactions

Eva Wiese
eva.wiese@tu-berlin.de
Technical University Berlin
Berlin, Germany

Nele Russwinkel
nele.russwinkel@uni-luebeck.de
University of Lübeck
Lübeck, Germany

ABSTRACT

Robots are more and more integrated into everyday environments, which makes it essential to examine how to design them so that humans are motivated and capable to interact with them. An important challenge for robotics is to determine how to design robots that accurately infer the human interaction partner's goals, intentions, and emotional states, and are able to adapt to their actions in time and space. Certain aspects of this challenge can be addressed through appropriate design of robot appearance and behavior, and equipping robots with appropriate models of social cognition. Other aspects, however, arise on the human side of the "equation", where lifelong experience with human interaction partners raises certain expectations of how verbal and nonverbal social cues are supposed to be interpreted, how actions are supposed to be coordinated and how emotional and motivational states are supposed to be communicated. If a robot meets these expectations, humans can interact with it quite intuitively, make accurate predictions regarding its actions and intentions, and interpret its social signals with ease. The question is how robots can be equipped with representations to meet these expectations. Most robots, however, violate anthropomorphic expectations in terms of their appearance, behavior and cognition, which can negatively impact performance, affect and motivation in human-robot interaction. In this paper, we discuss how the interplay between robots' actual capabilities and human expectations regarding these capabilities imposes challenges specifically for the time- dependent aspects of social human-robot interactions.

KEYWORDS

Time, Timing, Human-Robot Interaction, Cognitive Models, Expectations

1 CHALLENGES REGARDING TEMPORAL ASPECTS OF HRI

Interactions between humans and embodied social robots (i.e., human-robot interaction, HRI) face several critical challenges in terms of time perception and timing. **Firstly**, robots typically operate relatively slowly to prevent harm. Furthermore, due to latencies, technical constraints, and differences in degrees of freedom between a human's and a robot's motor system, robots are often not able to execute actions (e.g., grasping) with comparable spatial (e.g., different arrangement of joints forces a robot to grasp an object at a different point than a human would for a hand-over task) and temporal (e.g., constant motion speed compared to a human-typical speed-acceleration pattern) characteristics as humans do. This causes significant issues when trying to coordinate actions in space and time as it negatively impacts lower-level (e.g., entrainment) and higher-level aspects (e.g., inferring goals) of action understanding, planning and execution, and prevents fluent interactions.

Secondly, humans often expect robots to behave similarly to human partners because they likely lack specific knowledge about how robots work so that the anthropomorphic model is the only available relatable mental model in a socially interactive context. Once activated, the anthropomorphic mental model is used to make predictions and come up with explanations for observed behaviors. When interacting with a new system, the advantage of anthropomorphism is that it gives access to a very rich, experience-generated database that can account for lots of social interactive contexts; the disadvantage is that robots often do not behave like humans and in consequence violate the predictions generated by anthropomorphic mental models. Violation of expectations is known to be associated with increased cognitive load and negative emotions, which in the long run might be more disadvantageous for human-robot interaction than a robot that activated a more mechanistic mental model. The risk of violating expectations is specifically high when considering the temporal aspects of human-robot interactions given that even small deviations from expected trigger-response patterns (e.g., laughing at someone's joke, changing facial expressions after experiencing something painful) are noticed and disturb the fluency of human-robot communication and joint performance.

Thirdly, the human partner requires feedback in order to develop and calibrate a robot specific mental model and to correctly perform actions together with a robot. If a robot fails to give feedback and makes mistakes, it can cause frustration in the human partner [33]. Delays in the robot's feedback in response to a human input can be detrimental to the development of trust, and make it difficult for humans to adapt to the robot's input. Furthermore, if feedback in joint actions is not timed appropriately, it may not become transparent, which input triggered which output and thus makes coordination more difficult.

2 HOW TO ADDRESS TIME-DEPENDENT ISSUES IN HUMAN ROBOT INTERACTION?

To build robots that are perceived as intentional agents, we need to ask whether it is even necessary that they accurately emulate human behavior or whether it is sufficient for them to just display certain aspects of human behavior that are most strongly associated with the perception of intentionality [38]. Given the technological limitations associated with trying to reproduce large brain networks in artificial agents, the goal needs to be the identification

Veröffentlicht durch die Gesellschaft für Informatik e.V. in P. Fröhlich und Vanessa Cobus (Hrsg.): Mensch und Computer 2023 – Workshopband, 03.-06. September 2023, Rapperswil

https://doi.org/10.18420/muc2023-mci-ws05-452

of a minimal set of features that reliably trigger mind attribution to non-human agents. Neuroscientists and psychologists need to identify these features and investigate their effects on attitudes and performance in human–robot interactions. Engineers and computer scientists can then help with designing the robot body structure in a way that faithfully implements this minimal set of behavioral parameters in term of kinematics, dynamics, electronics, and computation. Lastly, cognitive scientists need to work on how artificial mental models can be generated to enable robots to better relate to their human partners' expectations. Trying to build robots that are perceived as intentional agents can also help to elucidate whether the minimal set of parameters relates to a specific architecture and how tuning various parameters affects the way a robot is perceived during the interaction with a human partner.

3 DIRECTIONS TOWARDS UNDERSTANDING TEMPORAL ASPECTS IN HRI

3.1 Temporal aspects of joint actions

3.1.1 Empirical studies on joint actions in HRI. One key mechanism in social interactions is the ability to understand the actions of others, that is: being able to tell what **sort of action** is executed, and based on what kind of **intention**. Action understanding is based on shared representations that are activated when an action is executed and when a similar action is observed in others (resonance; e.g., [14]). Observing others' actions facilitates the execution of similar actions, and hinders the execution of different actions since both action observation and execution activate the same neural network [19]. Shared representations are essential for performing joint actions, where two or more individuals coordinate their actions in time and space to achieve a shared action goal [30].

It is essential to understand if these phenomena are specific to human agents or if robot agents can also activate the action-perception system (APS). Robots were initially not assumed to activate the APS as it is sensitive to biological motion and goal-directed behavior. In line with this assumption, early studies on action understanding in HRI were not able to show resonance for the observation of robot actions [19] or at least to a significantly smaller degree than for the observation of human agents [22–24]. Follow-up studies showed that resonance can be induced by robots but that its degree seems to depend on the robot's physical appearance [6, 21], motion kinematics [3], or the body visibility [5]. Resonance during interactions with robots can even reach levels comparable to human-human interactions but only when participants are instructed to pay attention to the actions [8, 15, 37] or when given additional time to familiarize themselves with the observed actions [26]. In contrast, beliefs regarding a robot's human-likeness do not have an impact on the presence of motor resonance [25].

These studies suggest that robots have the potential to activate the APS, at the very least in a reduced fashion but under certain conditions even to a similar degree as in human-human interactions. The degree of APS activation depends on physical factors, such as the appearance and the kinematic profile of a robot, as well as cognitive factors, such as one's motivation to infer a robot's goals or the level of expertise with a particular robot system. This means that low-level mechanisms of social cognition are not specifically sensitive to the identity of an interaction partner, and can be activated by robots as long as their actions map onto the human motor repertoire, and people are motivated to pay attention to them.

3.1.2 Cognitive Modelling approach Action-Perception System. With cognitive modelling approaches such as with cognitive architectures e.g. ACT-R [1] it is possible to build flexible task models that can react to events (visual or auditory) and make decisions based on perceived information and pre-knowledge in a cognitive plausible way. Cognitive architectures refer to both a theory about the structure of the human mind and mental representations as well as a computational substantiation of such a theory. These architectures hold partly symbolic information and are therefor traceable in their behavior. There have been approaches to integrate subjective timing experience into this architectures, e.g. for accounting for differences in time perception due to experiences workload in complex tasks [28]. A model that uses such a cognitive architecture needs to have some understanding of the task (holds procedural knowledge), needs a goal (as a symbolic representation), and can build up a mental representation about the developing situation and the state of the human partner. An observer model, in contrast to just simulating a human doing a task, could build up representations observing the interaction situation. This provides the robot with a shared representations of task state and perceived information. This would enable the robot to behave more naturally since it can relate to simplified shared representations. There are examples of such models used for intelligent assistance for pilots understanding difficult situations and providing appropriate support [20]. This way delayed responses by pilots can be detected, and due to task knowledge and further information the cause of delay can be understood. Relating to the cause of delay appropriate steps can be initiated. For situations that require not just understanding of a delay or underlying intentions but also coordination of motor actions this is far more complex to realize. To enable joint-actions of human and robot different kind of models need to be integrated. First concepts of such an integrated approach that not just anticipates an action-intention on a cognitive level but also anticipates motor actions of the human and the planning motor actions of the robot need complex architectures [16]. After this has been implemented the temporal dependencies of such a joint are still an aspects that needs to be carefully tested.

3.2 Temporal aspects of understanding internal states

When navigating social environments, we need to understand how others feel (empathizing; [2]), and what they intend to do (mentalizing: [13]). Similar to joint action, empathizing and mentalizing are based on shared representations allowing us to infer the emotions and intentions of others by simulating what we would feel or intend in a comparable situation (e.g., [9]).

3.2.1 Empirical studies on understanding internal states. Perceiving emotional states in others activates similar emotional states in the observer, thereby creating shared representations at the neural and physiological level [27]. For instance, receiving a painful stimulus and observing the stimulus being presented to others activates similar brain networks [31]. Activation of similar brain networks has also been shown for both the presentation and observation of aversive olfactory [34] or haptic [18] stimuli. When studying mentalizing, researchers typically present participants with stories that involve false-belief manipulations [36] and require them to take the perspective of others in order to understand whether and how their representation of the situation differs from their own [10], make inferences about what others are interested in based on nonverbal cues like gaze direction [11], and reason about emotional states based on facial expressions or postures [2]. When we make inferences about the internal states of others, we need to incorporate knowledge about their dispositions and preferences into the mentalizing process [32]. This requires the ability to represent behaviors over a long period of time, across different circumstances and with different social partners [12].

These studies show that activation in brain areas related to empathizing and mentalizing are modulated by the degree to which interaction partners are perceived to have a mind, with stronger activation for intentional agents (i.e., humans) compared to non-intentional agents (i.e., robots; [35], for a review). Although further studies are necessary to determine the constraints under which robot agents activate the empathizing and mentalizing networks, the aforementioned studies provide evidence that activation in social brain areas involved in higher-order social-cognitive processes like empathizing and mentalizing more strongly depends on mind attribution than activation in social brain areas involved in lower-level social cognitive processing like action understanding.

3.2.2 Models of intention recognition. The recognition of intentions is relevant to adapt to others, especially in order to time actions in coordination with the other person. Several sources of information can help to successfully detect intentions of the partner. To understand intentions or to interact with others also simple concepts can help because humans can also relate to small kids or animals and understand intentions (e.g. [29] or [7]), therefore simple solutions can also offer robust support for better concepts in Human-Robot-Interactions. In order to realize this for robots the tracking of gaze can be used as in [17]. Here participants needed to choose different ingredients for a smoothie and the robot was able to identify the gaze direction and proactively moved forward towards the goal. This provided quicker robot actions and was perceived as helpful by the participants. On the other hand some people also rated this as weird because this behavior was not expected. Also more complex modelling approaches tried to anticipate human partners e.g. [4]. To use several sources of information to also identify intentions in cases of ambiguity or for mor complex tasks are interesting research challenges.

4 CHALLENGES FOR THE ROBOT INTERACTION PARTNER

In this paper it has been shown, that different aspects of timing in human-robot-interaction have a strong influence on the success of the interaction and also how this gives insight into cognitive processes that are involved. Different lines of research can shed light on the related mechanisms that are crucial here. The main challenges might guide interdisciplinary research done in this field. The first challenge was that robots operate relatively slowly due to technical reasons or due to safety reasons. To have robots that can react proactively or considers proactively dangerous upcoming situations would be beneficial. In order to build such system it is necessary to detect high-level intentions of the human partner. The second challenge focused on the issue that robots activate mental models relating to humans which often leads to violation of expectations. Therefore the question is, how can we design robot interactions that enables us to adapt our mental model well to the capabilities the robot has. The third challenge focused on the research question how to provide adequate feedback of the robot to the human, which includes challenge one and two. Research lines were shown how empirical studies and cognitive models provide possibilities to shed light on the mentioned issues. We hope that more exchange between research domains and approaches can be initiated by focusing on such challenges that are not domain specific.

REFERENCES

[1] John R. Anderson, Daniel Bothell, Michael D. Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. 2004. An integrated theory of the mind. *Psychological review* 111 4 (2004), 1036–60. https://doi.org/10.1037/0033-295X.111.4.1036

[2] Simon Baron-Cohen. 2005. The empathizing system. *Origins of the social mind: Evolutionary psychology and child development* (2005), 468–492.

[3] Ambra Bisio, Alessandra Sciutti, Francesco Nori, Giorgio Metta, Luciano Fadiga, Giulio Sandini, and Thierry Pozzo. 2014. Motor contagion during human-human and human-robot interaction. *PloS one* 9, 8 (2014), e106172.

[4] Sebastian Blum, Oliver Klaproth, and Nele Russwinkel. 2022. Cognitive Modeling of Anticipation: Unsupervised Learning and Sym-bolic Modeling of Pilots' Mental Representations. *Topics in Cognitive Science* 14, 4 (2022), 718–738.

[5] Thierry Chaminade and Gordon Cheng. 2009. Social cognitive neuroscience and humanoid robotics. *Journal of Physiology-Paris* 103, 3-5 (2009), 286–295.

[6] Thierry Chaminade, Jessica Hodgins, and Mitsuo Kawato. 2007. Anthropomorphism influences perception of computer-animated characters' actions. *Social cognitive and affective neuroscience* 2, 3 (2007), 206–216.

[7] Emily Collins. 2019. Drawing parallels in human-other interactions: a trans-disciplinary approach to developing human-robot interaction methodologies. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 374 (03 2019). https://doi.org/10.1098/rstb.2018.0433

[8] Emily S Cross, Roman Liepelt, Antonia F de C. Hamilton, Jim Parkinson, Richard Ramsey, Waltraud Stadler, and Wolfgang Prinz. 2012. Robotic movement preferentially engages the action observation network. *Human brain mapping* 33, 9 (2012), 2238–2254.

[9] Marie de Guzman, Geoffrey Bird, Michael J Banissy, and Caroline Catmur. 2016. Self–other control processes in social cognition: from imitation to empathy. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371, 1686 (2016), 20150079.

[10] Nicholas Epley, Carey K Morewedge, and Boaz Keysar. 2004. Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of experimental social psychology* 40, 6 (2004), 760–768.

[11] Chris Kelland Friesen and Alan Kingstone. 1998. The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review* 5, 3 (1998), 490–495.

[12] Uta Frith and Chris Frith. 2001. The biological basis of social interaction. *Current directions in psychological science* 10, 5 (2001), 151–155.

[13] Uta Frith and Christopher D Frith. 2003. Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358, 1431 (2003), 459–473.

[14] Vittorio Gallese, Luciano Fadiga, Leonardo Fogassi, and Giacomo Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain* 119, 2 (04 1996), 593–609. https://doi.org/10.1093/brain/119.2.593 arXiv:https://academic.oup.com/brain/article-pdf/119/2/593/1084858/119-2-593.pdf

[15] Valeria Gazzola, Giacomo Rizzolatti, Bruno Wicker, and Christian Keysers. 2007. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35, 4 (2007), 1674–1684.

[16] Chenxu Hao, Nele Russwinkel, Daniel F.B. Haeufle, and Philipp Beckerle. 2023. A Commentary on Towards autonomous artificial agents with an active self: Modeling sense of control in situated action. *Cognitive Systems Research* 79 (2023), 1–3. https://doi.org/10.1016/j.cogsys.2022.12.006

[17] Chien-Ming Huang and Bilge Mutlu. 2016. Anticipatory robot control for efficient human-robot collaboration. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 83–90. https://doi.org/10.1109/HRI.2016.7451737

[18] Christian Keysers and David I Perrett. 2004. Hebbian learning. *Trends in Cognitive Sciences* 11, 8 (2004), 501–507.

[19] J.M Kilner, Y Paulignan, and S.J Blakemore. 2003. An Interference Effect of Observed Biological Movement on Action. *Current Biology* 13, 6 (2003), 522–525. https://doi.org/10.1016/S0960-9822(03)00165-9

[20] Oliver W Klaproth, Marc Halbrügge, Laurens R Krol, Christoph Vernaleken, Thorsten O Zander, and Nele Russwinkel. 2020. A neuroadaptive cognitive model for dealing with uncertainty in tracing pilots' cognitive state. *Topics in cognitive science* 12, 3 (2020), 1012–1029.

[21] Aleksandra Kupferberg, Markus Huber, Bartosz Helfer, Claus Lenz, Alois Knoll, and Stefan Glasauer. 2012. Moving just like you: motor interference depends on similar motility of agent and observer. *PloS one* 7, 6 (2012), e39637.

[22] Lindsay M Oberman, Joseph P McCleery, Vilayanur S Ramachandran, and Jaime A Pineda. 2007. EEG evidence for mirror neuron activity during the observation of human and robot actions: Toward an analysis of the human qualities of interactive robots. *Neurocomputing* 70, 13-15 (2007), 2194–2203.

[23] Erhan Oztop, David W Franklin, Thierry Chaminade, and Gordon Cheng. 2005. Human–humanoid interaction: is a humanoid robot perceived as a human? *International Journal of Humanoid Robotics* 2, 04 (2005), 537–559.

[24] Clare Press, Geoffrey Bird, Rüdiger Flach, and Cecilia Heyes. 2005. Robotic movement elicits automatic imitation. *Cognitive brain research* 25, 3 (2005), 632–640.

[25] Clare Press, Helge Gillmeister, and Cecilia Heyes. 2006. Bottom-up, not top-down, modulation of imitation by human and robotic models. *European Journal of Neuroscience* 24, 8 (2006), 2415–2419.

[26] Clare Press, Helge Gillmeister, and Cecilia Heyes. 2007. Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of the Royal Society B: Biological Sciences* 274, 1625 (2007), 2509–2514.

[27] Stephanie D. Preston and Frans B. M. de Waal. 2002. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences* 25, 1 (2002), 1–20. https://doi.org/10.1017/S0140525X02000018

[28] Nele Russwinkel, Leon Urbas, and Manfred Thüring. 2011. Predicting temporal errors in complex task environments: A computational and experimental approach. *Cognitive Systems Research* 12, 3 (2011), 336–354. https://doi.org/10.1016/j.cogsys.2010.09.003 Special Issue on Complex Cognition.

[29] Enrique Schaerer, Richard Kelley, and Monica Nicolescu. 2009. Robots as animals: A framework for liability and responsibility in human-robot interactions. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*. 72–77. https://doi.org/10.1109/ROMAN.2009.5326244

[30] Natalie Sebanz and Guenther Knoblich. 2009. Prediction in joint action: What, when, and where. *Topics in cognitive science* 1, 2 (2009), 353–367.

[31] Tania Singer, Ben Seymour, John O'doherty, Holger Kaube, Raymond J Dolan, and Chris D Frith. 2004. Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 5661 (2004), 1157–1162.

[32] Frank Van Overwalle. 2009. Social cognition and the brain: a meta-analysis. *Human brain mapping* 30, 3 (2009), 829–858.

[33] Alexandra Weidemann and Nele Rußwinkel. 2021. The Role of Frustration in Human–Robot Interaction–What Is Needed for a Successful Collaboration? *Frontiers in Psychology* (2021), 707. https://doi.org/10.3389/fpsyg.2021.640186

[34] B Wicker et al. [n.d.]. Keysers, Christian; Plailly, Jane; Royet, Jean-Pierre; Gallese, Vittorio; Rizzolatti, Giacomo (2003)." Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 3 ([n. d.]), 655–664.

[35] Eva Wiese, Giorgio Metta, and Agnieszka Wykowska. 2017. Robots as intentional agents: using neuroscientific methods to make robots appear more social. *Frontiers in psychology* 8 (2017), 1663.

[36] Heinz Wimmer and Josef Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 1 (1983), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

[37] Agnieszka Wykowska, Ryad Chellali, Md Mamun Al-Amin, and Hermann J Müller. 2014. Implications of robot actions for human perception. How do we represent actions of the observed robots? *International Journal of Social Robotics* 6 (2014), 357–366.

[38] Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2006. How contingent should a communication robot be?. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 313–320.