Proactive eye-gaze in human-robot interaction[1]

Erik Billing[1], Alessandra Sciutti[2], and Giulio Sandini[2]

[1]University of Skövde, Skövde, Sweden

[2]Italian Institute of Technology, Genova, Italy

February 2019

1 Introduction

Robot technology has over the last decades developed into many new areas and applications. From being a technology primarily associated with industrial automation where robots are separated from people using e.g., safety cages, we now see a flourish of new applications where robots are designed to interact with humans. Robots are serving at restaurants (Webster, 2018), acting as companions for elderly (PARO, 2019), and constituting interaction partners for many types of games and playful applications (Anki, 2019). In the industrial domain we see a strong trend towards *human-robot collaboration (HRC)*, with robots like

---

[1] The work has been financially supported by the Knowledge Foundation, Stockholm, under SIDUS grant agreement no. 20140220 (AIR, “Action and intention recognition in human interaction with autonomous systems”).

Sawyer by Rethink Robotics and ABB YuMi. These robots are designed to interact with people in that they are safe, but are not yet particularly collaborative. In some respects these robot are working close to humans and not necessarily "with" humans.

Indeed one the main missing skill in current machines is the inability to anticipate and predict the human partners' behaviors (Sandini and Sciutti, 2018). Conversely, humans are always projected into the future, continuously imagining their actions and their potential effects through simulations mediated by internal models (Bhat et al., 2016). As a result, when observing someone's else action, we are already predicting its consequences and the actor's goal, since infancy (Meltzoff, 1995). Collaboration and coordination between humans cannot be easily achieved without such prospective ability and there are many reasons to believe that the anticipatory nature of collaboration applies equally in *human-robot interaction (HRI)* as it does between humans (Vernon et al., 2016).

Some evidence of the advantages of introducing anticipation into HRI can be found in the literature. Hoffman (2010) studied the effects of reactive vs anticipatory robot control and found positive effects of anticipatory control on perceived fluency of interaction, and in one case, on team efficiency. Huang and Mutlu (2016) make use of eye-tracking to predict the actions of a human user in a pick and place scenario, showing that the task can be executed faster when the robot is anticipating the user's actions. Mainprice and Berenson (2013) evaluate the effects of early prediction of human motion during human-robot collaboration, showing that the robot is able to safely avoid the human even when initial predictions of the human's motion are incorrect. Taking the reversed perspective, it is also crucial for a robot to intuitively communicate its goals to its human partners in order to allow for anticipation from their side (Sciutti et al., 2018). In several studies focused on human-robot collaboration, it has been proven that a highly "readable" (or legible) robot motion leads to a positive evaluation of the robot and to an increase in efficiency of the interaction (Dragan et al., 2015). Following a different approach, Chadalavada et al. (2015) evaluated a method for communicating robot intentions using projections of future movement trajectories. Positive effects on user ratings of the robot's communication, reliability, predictability, transparency and situation awareness were found. The authors highlight that already simple information, such as the trajectory projections, can effectively improve user experience. In a similar vein, Watanabe et al. (2015) presents a method for intention communication for a robotic wheelchair, using light projection Evaluation results show preferences for navigational intention communication, both for the wheelchair passenger and persons passing by.

While these studies effectively demonstrate the value of anticipation in interaction, still it is to be fully understood how such anticipation is achieved by humans. Humans' ability to anticipate the actions of others is believed to stem from the *mirror-neuron system (MNS)* and provides a direct matching of observed actions onto the observer's own motor system (Flanagan and Johansson, 2003). Specifically, the motor system of the observer is activated during action observation and appears to resonate with that of the actor (Rizzolatti et al., 1999). Exactly which circumstances that trigger direct matching is still largely unknown. A better understanding of the neurological basis for action execution and observation could provide valuable insights for HRI (Sciutti et al., 2012a). With a rapidly growing body of research studying people's perceptions of robots, ranging from the much debated *Uncanny Valley* (Mori, 1970) to the use of standardized questionnaires such as the *negative attitudes toward robots scale* (Nomura et al., 2006), we believe that it is crucial to complement these studies with explanations of the cognitive processes underlying the perception and understanding of robots.

One of the most common ways to study action anticipation is the analysis of agents' gaze to identify the presence of proactive eye-gaze (PEG). When humans manipulate objects, they typically fixate at the goal of the action, producing a eye-gaze pattern that is proactive in relation to the hand (Johansson et al., 2001). This proactive eye-gaze coordination is mirrored by the observer, even when the eyes of the actor is not visible (Flanagan and Johansson, 2003). This phenomenon is believed to stem from recruitment of the MNS, and thus, from a direct matching of observed actions onto the motor system of the observer.

While a huge body of literature is building up on the emergence of proactive gaze, it is almost exclusively concerned with human or animal actors, leaving PEG during observation of robots relatively unstudied. In the only counter-example that we are aware of, Sciutti et al. (2012b) demonstrated that robot actions can, under specific circumstances, trigger PEG, most likely resulting from MNS activation and direct matching of the observed actions. This result opens new possibilities to design robots so that they can resonate with the motor system of their human users. Still it is to be clearly understood which elements are necessary to allow for the emergence of automatic robot action anticipation, triggering PEG.

With the ambition of designing robots that can make it easier for humans to predict the robot's actions, by eliciting motor resonance and

PEG among their users, we here propose three open questions linking proactive gaze and HRI:

1. Which aspects of observed action triggers action anticipation with the gaze?
2. Under which conditions does MNS activation lead to PEG or other observable cues, supporting collaboration and joint action?
3. To what extent does MNS activation correlate with improvements in human-robot collaboration?

Concerning question 1, MNS activation has been linked to presence of biological motion (Saygin et al., 2004; Ulloa and Pineda, 2007). Elsner et al. (2012) used a point-light display of reaching actions and demonstrated that proactive gaze appears when the hand follows a standard, biological motion profile, but not when the acceleration pattern was manipulated to a linear (mechanical motion) form. Elsner et al. concludes that *kinematic information from biological motion can be used to anticipate the goal of other people's point-light actions and that the presence of biological motion is sufficient for anticipation to occur.*

This explanation may however not be conclusive. Already the initial study by Flanagan and Johansson (2003) comprised a self-propelled condition that did not elicit proactive gaze, despite the fact that the object moved with a biological motion profile. Additionally, using fMRI, Gazzola et al. (2007) found similar MNS activation during observation of both human and robot actions, despite the fact that robot motion was clearly non-biological.

Binding to question 2, Ambrosini et al. (2011) investigated proactive gaze behavior during reaching actions towards multiple targets, using either a hand pre-shaped towards grasping one of the targets or a closed fist. They found PEG only when the reaching hand adopted a grasping preshape, but not when the hand was closed. Both conditions adopted a biological motion profile and are likely to activate the MNS. Thus, the lack of PEG in the control condition could be understood as an a result of a more complex association between MNS and PEG than we've previously been aware of.

Finally, question 3 concerns the link between MNS activation and concrete benefits for collaboration. While the MNS is commonly described as *the* link between action observation and execution, its necessary involvement is not fully clarified. Gredebäck and Melinder (2010) investigated action anticipation in 6 and 12 month old infants observing feeding actions. PEG was observed among 12 month old subject, but not in the younger infants. However, both groups demonstrated pupil dilation in response to non-rational actions, suggesting that also the 6 month old infants can interpret the goal of observed action without MNS recruitment. Gredebäck and Melinder suggest a dual-route explanations for their results, raising the question to what extent also adults can rely on a second route to action anticipation in cases when the MNS is not activated.

In conclusion, when designing robots for collaboration with humans, precise timing of actions is critical for many applications. A mutual ability to anticipate the actions of the other would allow robots to adapt to the user in a way that is not happening today, also increasing safety. PEG provides a potentially very useful cue, both for anticipating the users' actions, and for communicating planned robot actions in a way that is automatically interpreted by the human user. Therefore we claim that in the next future it is worth investigating in depth which factors influence PEG during Human-Robot Interaction, so as to facilitate mutual anticipation during collaboration with machines.

References

E. Ambrosini, M. Costantini, and C. Sinigaglia. Grasping with the eyes. *Journal of Neurophysiology*, 106 (3): 1437–1442, sep 2011. ISSN 0022-3077. 10.1152/jn.00118.2011.

Anki. Anki — We create robots that move you, 2019. URL `https://www.anki.com/en-us`.

A. A. Bhat, V. Mohan, G. Sandini, and P. Morasso. Humanoid infers Archimedes' principle: Understanding physical relations and object affordances through cumulative learning experiences. *Journal of the Royal Society Interface*, 13 (120), 2016. ISSN 17425662. 10.1098/rsif.2016.0310.

R. T. Chadalavada, H. Andreasson, R. Krug, and A. J. Lilienthal. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *2015 European Conference on Mobile Robots (ECMR)*, pages 1–6. IEEE, sep 2015. ISBN 978-1-4673-9163-4. 10.1109/ECMR.2015.7403771.

A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa. Effects of Robot Motion on Human-Robot Collaboration. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15*, pages 51–58, New York, New York, USA, 2015. ACM Press. ISBN 9781450328838. 10.1145/2696454.2696473.

C. Elsner, T. Falck-Ytter, and G. Gredebäck. Humans Anticipate the Goal of other People's Point-Light Actions. *Frontiers in psychology*, 3: 120, 2012. ISSN 1664-1078. 10.3389/fpsyg.2012.00120.

J. R. Flanagan and R. S. Johansson. Action plans used in action observation. *Nature*, 424 (6950): 769–771, 2003. ISSN 0028-0836. 10.1038/nature01861.

V. Gazzola, G. Rizzolatti, B. Wicker, and C. Keysers. The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage*, 35 (4): 1674–1684, 2007. ISSN 10538119. 10.1016/j.neuroimage.2007.02.003.

G. Gredebäck and A. Melinder. Infants' understanding of everyday social interactions: A dual process account. *Cognition*, 114 (2): 197–206, 2010. ISSN 00100277. 10.1016/j.cognition.2009.09.004.

G. Hoffman. Anticipation in Human-Robot Interaction. *2010 AAAI Spring Symposium Series*, pages 21–26, 2010.

C. M. Huang and B. Mutlu. Anticipatory robot control for efficient human-robot collaboration. *ACM/IEEE International Conference on Human-Robot Interaction*, pages 83–90, 2016. ISSN 21672148. 10.1109/HRI.2016.7451737.

R. S. Johansson, G. Westling, A. Bäckström, and J. R. Flanagan. Eye–Hand Coordination in Object Manipulation. *Journal of Neuroscience*, 21 (17): 6917–6932, sep 2001. ISSN 0270-6474. 10.1523/JNEUROSCI.21-17-06917.2001.

J. Mainprice and D. Berenson. Human-robot collaborative manipulation planning using early prediction of human motion. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 299–306. IEEE, nov 2013. ISBN 978-1-4673-6358-7. 10.1109/IROS.2013.6696368.

A. N. Meltzoff. Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental psychology*, 31 (5): 838–850, sep 1995. ISSN 1939-0599. 10.1037/0012-1649.31.5.838.

M. Mori. Bukimi no tani (The uncanny valley). *Energy*, 7 (4): 33–35, 1970.

T. Nomura, T. Suzuki, T. Kanda, and K. Kato. Measurement of negative attitudes toward robots. *Interaction Studies*, 7 (3): 437–454, nov 2006. ISSN 1572-0373. 10.1075/is.7.3.14nom.

PARO. PARO Therapeutic Robot, 2019. URL `http://www.parorobots.com/`.

G. Rizzolatti, L. Fadiga, L. Fogassi, and V. Gallese. Resonance Behaviors and Mirror Neurons. *Archives Italiennes de Biologie*, 137 (2): 85–100, may 1999. ISSN 0003-9829. 10.4449/AIB.V137I2.575.

G. Sandini and A. Sciutti. Humane Robots—from Robots with a Humanoid Body to Robots with an Anthropomorphic Mind. *ACM Transactions on Human-Robot Interaction (THRI)*, 7 (1): 7, 2018. 10.1145/3208954.

A. P. Saygin, S. M. Wilson, D. J. Hagler, E. Bates, and M. I. Sereno. Point-Light Biological Motion Perception Activates Human Premotor Cortex. *Journal of Neuroscience*, 24 (27): 6181–6188, jul 2004. ISSN 0270-6474. 10.1523/JNEUROSCI.0504-04.2004.

A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, T. Pozzo, and G. Sandini. Measuring Human-Robot Interaction Through Motor Resonance. *International Journal of Social Robotics*, 4 (3): 223–234, aug 2012a. ISSN 1875-4791. 10.1007/s12369-012-0143-1.

A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, and G. Sandini. Anticipatory gaze in human-robot interactions. In *Proceedings of the Gaze in Human-Robot Interaction Workshop held at the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2012).*, number March, 2012b.

A. Sciutti, M. Mara, V. Tagliasco, and G. Sandini. Humanizing human-robot interaction: On the importance of mutual understanding. *IEEE Technology and Society Magazine*, 37 (1): 22–29, mar 2018. ISSN 02780097. 10.1109/MTS.2018.2795095.

E. R. Ulloa and J. A. Pineda. Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behavioural Brain Research*, 183 (2): 188–194, 2007. ISSN 01664328. 10.1016/j.bbr.2007.06.007.

D. Vernon, S. Thill, and T. Ziemke. The Role of Intention in Cognitive Robotics. In A. Esposito and L. C. Jain, editors, *Toward Robotic Socially*

Believable Behaving Systems, pages 15–27. Springer, 2016. ISBN 9789264154643. 10.1787/9789264192263-en.

A. Watanabe, T. Ikeda, Y. Morales, K. Shinozawa, T. Miyashita, and N. Hagita. Communicating robotic navigational intentions. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5763–5769. IEEE, sep 2015. ISBN 978-1-4799-9994-1. 10.1109/IROS.2015.7354195.

N. Webster. Ruby the robot waitress serves up a treat in Dubai. *The National*, jul 2018. URL `https://www.thenational.ae/uae/ruby-the-robot-waitress-serves-up-a-treat-in-dubai-1.751222`.