Distinguishing minds in interaction: Modeling self-other distinction in the motor system

Sebastian Kahl and Stefan Kopp

Social Cognitive Systems Group, Faculty of Technology,
Bielefeld University, Inspiration 1, 33619 Bielefeld, Germany
{skahl,skopp}@uni-bielefeld.de

In order to engage in a continuous social interaction, participants must be able to dynamically understand, predict, and influence each other's mental states and actions, so as to enable a process of efficient and interactive grounding of shared meaning. We follow the argument that the mentalizing network and the mirror-neuron system in our social brain together provide the basis for these abilities [1]. However, how these systems exactly operate and how they work together is still unclear. Building on previous work on the interplay of mentalizing and mirroring in embodied communication, here we lay out our modeling approach together with simulation results that show how to solve the problem of self-other distinction in a model of the sensorimotor system of interacting agents. We believe this to be an important step that informs mentalizing, so that in the future virtual agents equipped with this model can distinguish between own and other's mental states in complex social situations.

In a first step towards this goal, we developed a model of two distinct networks of the human social brain - mentalizing and mirroring - which allows them to interact during embodied communication. The model connects a mentalizing system based on simple heuristics for attributing and inferring different orders of belief about own and other's mental states, with a hierarchical predictive processing model of online action perception and production based on the common coding of underlying action representations [2]. To investigate the role of mentalizing and mirroring interacting in inter-agent coordination and to test the model, we conducted simulation experiments in which two virtual agents were each equipped with this model. Different mentalizing capacity configurations were tested, as well as different noise conditions, thus influencing the robustness of the communication. The agents engage in non-verbal communication behavior to which the embodied action representations in the mirroring system can resonate because of their close coupling of perception and production, while taking uncertainty from noise into account. Resonating action representations inform the mentalizing system, which in turn can guide successful interaction. Results from our simulations on this first model demonstrate how mentalizing can afford higher robustness of communication by enabling interactive grounding processes.

Although our model was able to act upon and infer beliefs about own and other's mental states, it could only produce or perceive an action at a time. Of course, this is a special case of interaction that can occur, but in our dynamic world our social brain has to cope with concurrent production and perception situations, e.g., dancing with a partner or playing a duet on the piano. Situations where you have to simultaneously keep track of your own actions while coordinating with your interaction partner's actions, all using the same sensorimotor processes. Thus, as a starting point for an account of simultaneous action and perception, the first step is to enable early self-other distinction within the sensorimotor system, which we will use to distinguish ourselves from others so that we do not falsely attribute an action outcome to ourselves. As research into schizophrenia has shown, reliable and early self-other integration and distinction is important not only for the correct attribution of a sense of agency, but also for the correct attribution of intentions and emotions in social interactions [3]. We modeled two processes that have been identified in the literature, which together can be integrated into a combined sense of agency. The **predictive process** makes use of people's ability to anticipate the sensory consequences of their own actions. One account to model these processes is based on inverse and forward models [4]. The second, **postdictive process** uses higher-level causal beliefs like the intention to act [5] and temporal binding as mechanisms to infer the consistency of the action outcome. The "temporal binding window" [6] describes the integration of sensory signals from multiple modalities, which selects perceived actions and their outcomes for integration as long as they occur within a narrow temporal window. Because we have more experience in predicting our own body that window is more narrow for own action-outcomes than for other people's actions. Being able to make such a distinction allows people to monitor, infer and distinguish between causal relations for own and other's behavior.

But how do these two processes work together to inform a sense of agency and what if cues for agency are unreliable? We extended our prediction-based model of the sensorimotor system with a model of self-other distinction based on postdictive and predictive cues and compared two ways of cue integration to the literature on the sense of agency. In the predictive process, we modeled the match or mismatch of the predicted action-outcome. In the postdictive process, we modeled the intention to act and the delay in the action-outcome for temporal binding (see [7] for more detail). It is not obvious from the literature how the cues from the predictive and postdictive processes are being integrated. As a first step, we test two simple ways to do so, namely, to connect them conjunctively or disjunctively. A conjunctive connection would allow a sense of agency to occur only if it is supported by both processes; in a disjunctive connection only one cue would suffice to inform a sense of agency, in a more flexible but also more error prone manner.

To test the extended model's ability to differentiate between self and other we simulated a social situation. In this situation, the model writes a character from a learned set of handwritten characters of the alphabet while it either receives the correct action-outcome as feedback, or it receives delayed or different feedback than expected. This way we simulate the effect of simultaneously perceiving an interaction partner's action. We tested three scenarios for each combination to logically connect predictive and postdictive cues to form a sense of agency. In the first scenario we trigger the intent to act out a motor schema and the model will perceive its own correct output as feedback. Here, the model receives both cues correctly. In the second scenario we trigger the same intent to act, but the model now receives its output with a delay, disrupting the postdictive cue. In the third scenario the model again is triggered to act, while this time it receives unpredicted action-outcomes.

Results from these simulations show that a disjunctive connection between cues fares better, since it allows for a sense of agency even in dynamic scenarios of simultaneous perception and production, with feedback either from own or from other's actions. This more flexible distinction is also supported by results reported in the literature, e.g., [8] found that where the reliability of the predictive process was reduced, the system put more weight on the post-hoc inferential processes. Taken together, our modeling approach supports the motor system's role in social cognition. In future work we want to use the differentiating information provided by the present model to inform higher-level cognition through an interplay with the mentalizing system, to prepare virtual agents to distinguish between own and other's mental states in complex social situations.

Acknowledgements This research/work was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

References

1. Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., Vogeley, K.: Toward a second-person neuroscience. Behav Brain Sci. 36, 393414 (2013).
2. Kahl, S., Kopp, S.: Towards a Model of the Interplay of Mentalizing and Mirroring in Embodied Communication. In: Airenti, G., Bara, B.G., and Sandini, G. (eds.) Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science. 1419, p. 300-305. (2015).
3. van der Weiden, A., Prikken, M., van Haren, N.E.M.: Self-other integration and distinction in schizophrenia: A theoretical analysis and a review of the evidence. Neuroscience and Biobehavioral Reviews. 57, 220237 (2015).
4. Frith, C. D., Blakemore, & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. Philosophical Transactions of the Royal Society B: Biological Sciences, 355(1404), 17711788.
5. Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation. Sources of the experience of will. American Psychologist, 54(7), 480492.
6. Colonius, H., Diederich, A.: Multisensory Interaction in Saccadic Reaction Time: A Time-Window-of-Integration Model. Journal of Cognitive Neuroscience. 16, 10001009 (2004).
7. Kahl, S., & Kopp, S. (Accepted). Self-other distinction in the motor system during social interaction: A computational model based on predictive processing. Presented at the 39th Annual Meeting of the Cognitive Science Society (CogSci 2017).
8. Wolpe, N., Moore, J. W., Rae, C. L., Rittman, T., Altena, E., Haggard, P., & Rowe, J. B. (2014). The medial frontal-prefrontal network for altered awareness and control of action in corticobasal syndrome. Brain, 137(1), 208220.