Modeling the integration of active inference and sense of agency for self-other distinction

Sebastian Kahl and Stefan Kopp

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

People are able to infer a sense of agency for their actions, i.e., the sense that an action is self-generated. We hypothesize that a sense of agency also allows to infer an early (motor coordination level) distinction of own from other's actions during non-verbal social interaction. To test this, we simulated the handwriting of numerals using a computational model based on principles of active inference. The model has no sense of another agent being present and it cannot see its own hand and the pen it holds. It only sees what is written on a white canvas, and has the ability to write. This is all the model knows about its world. Now when it writes, all it can do to distinguish its own action from that of another agent, is to recognize its own writing. This is especially tricky when another agent would try to write at the same time. We have run several simulation scenarios and will show the dynamics of the model's behavior and judgements of sense of agency [1].

We developed and implemented a generative model of the dynamic processes underlying motor coordination and self-other distinction. It consists of a hierarchy of predictive processes of increasingly abstract representations over visuomotor primitives. These representations are also generative processes, which together form a hierarchical generative model which maps from (hidden) causes in the world to their perceived (sensory) consequences. The model relies on principles of active inference and free energy minimization [2, 3]. The tight coupling between action and perception in active inference means that, following prediction errors, either the model hypotheses have to be updated or action in the world is necessary to make future sensory evidence meet the model predictions. Free energy is merely the term for the negative log model evidence of a perceived event given the model prediction, i.e., the prediction error which is to be minimized.

The hierarchical model makes use of spatial and temporal aspects of action, and integrates predictive and postdictive cues for sense of agency over time. To be more specific, what allows the model to infer that it is writing a number itself, is not only the prediction of the actual stroke it intended to write, as in the comparator model [4]. In addition, the time it took from an intention to act until its perceived consequence needs to be within the predicted bounds of its embodiment [5, 6]. That is, it can predict not only how its body needs to act, but also how long it takes to act. This, along with a fluency effect and higher level intentions to act, allow the model to attribute a sense of agency to own actions.

We simulated and compared different scenarios of perception and production of handwritten numerals to showcase the belief dynamics, and how active inference works to minimize free energy on the different levels of the model's hierarchy. In addition, we show how temporal and spatial prediction-errors increase free energy, decrease precision, and influence the integration of sense of agency over time.

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. Kahl, S., & Kopp, S.: A Predictive Processing Model of Perception and Action for Self-Other Distinction. Frontiers in Psychology, 9, p. 47–14. (2018)
2. Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J.: Action and behavior: a free-energy formulation. Biol. Cybern. 102, 227–260. (2010)
3. Friston, K., Mattout, J., & Kilner, J.: Action understanding and active inference. Biol. Cybern. 104, 137–160. (2011)
4. Frith, C. D., Blakemore, S. J., & Wolpert, D. M.: Abnormalities in the awareness and control of action. Philos. Trans. R. Soc. B Biol. Sci. 355. (2000)
5. Colonius, H., & Diederich, A.: Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J. Cogn. Neurosci. 16, 1000–1009. (2004)
6. 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. Neurosci. Biobehav. Rev. 57, 220–237. (2015)