pp. S32–S33, Tübingen, Germany. doi:10.1007/s10339-014-0632-2
The final publication is available at link.springer.com

## A model for dynamic minimal mentalizing in dialogue

*Hendrik Buschmeier, Stefan Kopp*
*Social Cognitive Systems Group, CITEC and Faculty of Technology, Bielefeld University*

Spontaneous dialogue is a highly interactive endeavour in which interlocutors constantly influence each other's actions. As addressees they provide feedback of perception, understanding, acceptance, and attitude [1]. As speakers they adapt their speech to the perceived needs of the addressee, propose new terms and names, make creative references, draw upon established and known to be shared knowledge, etc. This makes dialogue a 'joint activity' [7] whose outcome is not determined up front but shaped by the interlocutors while the interaction unfolds over time.

One of the tasks interlocutors need to carry out while being engaged in a dialogue is keeping track of the dialogue information state. This is usually considered to be a rich representation of the dialogue context, most importantly including which information is grounded and which is still pending to be grounded (and potentially much more information; see, e.g., [10]). Whether such a detailed representation of the information state is necessary – and whether it is a cognitively plausible assumption – for participating in dialogue is a topic of ongoing debate.

On the one hand, Brennan and Clark [2,7] state that speakers maintain a detailed model of common ground and design their utterance to the exact needs of their communication partners – even to the extent that approximate versions of mutual knowledge may be necessary to explain certain dialogue phenomena [8]. On the other hand, Pickering and Garrod [11] argue that – for reasons of efficiency – dialogue cannot involve heavy inference on common ground, but is an automatic process that relies on priming and activation of linguistic representations and uses interactive repair upon miscommunication. A position that falls in between this dichotomy is Galati and Brennan's [9] lightweight one-bit partner model (e.g., has the addressee heard this before or not) that can be used instead of full common ground when producing an utterance.

We propose that interlocutors in dialogue engage in *dynamic minimal mentalizing*, a process that goes beyond the single properties in the focus of Galati and Brennan's [9] 'one-bit' model, but is comparable in computational efficiency. We assume that speakers maintain a probabilistic, multidimensional (consisting of a fixed number of state variables), and dynamic 'attributed listener state' [5]. We model this as a dynamic Bayesian network representation (see Fig. 1) that is continuously updated by the addressees' communicative feedback (i.e., short verbal-vocal expressions such a 'uh-huh,' 'yeah,' 'huh?'; head gestures; facial expressions) seen as evidence of understanding in response to ongoing utterances.

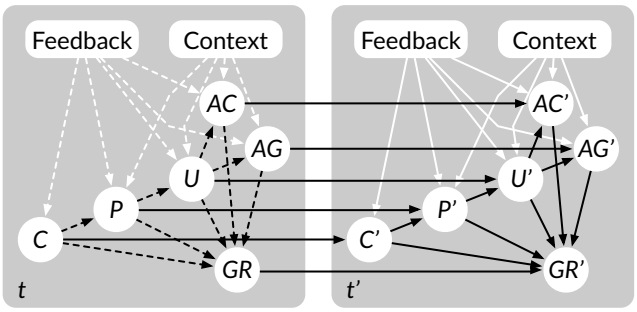


**Figure 1** The dynamic Bayesian network model for dynamic minimal mentalizing. The network consists of the mental state variables for contact (*C*), perception (*P*), understanding (*U*), acceptance (*AC*), agreement (*AG*), and groundedness (*GR*) attributed to the listener.

The proposed model is *multidimensional* because it represents the listeners' mental state of listening in terms of the various communicative functions that can be expressed in feedback [1]: is the listener in *contact?;* is he or she willing and able to *perceive* and *understand* what is said?; and does he or she *accept* the message and *agrees* to it? Instead of making a decision conditioned on the question whether the interlocutor has heard something before, this model allows to make use of the still computationally feasible but richer knowledge of whether he or she has likely perceived, understood, etc. a previously made utterance.

Further, the model is fully *probabilistic* since the attributed mental states are modelled in a Bayesian network. Each dimensions is represented as a random variable and the probabilities over the state of each variable (e.g., low, medium, high understanding) are interpreted in terms of the speaker's degree of belief in the addressee being in a specific state. This is a graded form of common ground [3] and presupposition (e.g., this knowledge is most likely in the common ground; see variables *GR* and *GR′* in Figure 1), which can be accommodated by, e.g., interactively leaving information out or adding redundant information; or by making information pragmatically implicit or explicit.

Finally, since the model is based on a *dynamic* Bayesian network, the interpretation of incoming feedback signals from the addressee is influenced by the current belief state, and changes of the attributed listener state are tracked over time. Representing these dynamics provides speakers with a broader basis for production choices as well as enabling strategic placement of feedback elicitation cues based on informational needs. It also allows for a prediction of the addressee's likely future mental state, thus enabling anticipatory adaptation of upcoming utterances.

In current work, the model of dynamic minimal mentalizing is being applied and evaluated in a virtual conversational agent that is able to interpret its user's communicative feedback and adapt its own language accordingly [4,6].

**Acknowledgements** This research is supported by the Deutsche Forschungsgemeinschaft (DFG) through the Center of Excellence EXC 277 'Cognitive Interaction Technology.'

## References

1. Allwood, J., Nivre, J., Ahlsén, E.: On the semantics and pragmatics of linguistic feedback. Journal of Semantics **9**, 1–26 (1992). DOI 10.1093/jos/9.1.1
2. Brennan, S.E., Clark, H.H.: Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory, and Cognition **22**, 1482–1493 (1996). DOI 10.1037/0278-7393.22.6.1482
3. Brown-Schmidt, S.: Beyond common and privileged: Gradient representations of common ground in real-time language use. Language and Cognitive Processes pp. 62–89 (2012). DOI 10.1080/01690965.2010.543363
4. Buschmeier, H., Kopp, S.: Towards conversational agents that attend to and adapt to communicative user feedback. In: Proceedings of the 11th International Conference on Intelligent Virtual Agents, pp. 169–182. Reykjavík, Iceland (2011). DOI 10.1007/978-3-642-23974-8_19
5. Buschmeier, H., Kopp, S.: Using a Bayesian model of the listener to unveil the dialogue information state. In: SemDial 2012: Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue, pp. 12–20. Paris, France (2012)
6. Buschmeier, H., Kopp, S.: When to elicit feedback in dialogue: Towards a model based on the information needs of speakers. In: Proceedings of the 14th International Conference on Intelligent Virtual Agents, pp. 71–80. Boston, MA, USA (2014)
7. Clark, H.H.: Using Language. Cambridge University Press, Cambridge, UK (1996). DOI 10.1017/CBO9780511620539
8. Clark, H.H., Marshall, C.R.: Definite reference and mutual knowledge. In: A.K. Joshi, B.L. Webber, I.A. Sag (eds.) Elements of Discourse Understanding, pp. 10–63. Cambridge University Press, Cambridge, UK (1981)
9. Galati, A., Brennan, S.E.: Attenuating information in spoken communication: For the speaker, or for the addressee? Journal of Memory and Language **62**, 35–51 (2010). DOI 10.1016/j.jml.2009.09.002
10. Ginzburg, J.: The Interactive Stance. Oxford University Press, Oxford, UK (2012)
11. Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences **27**, 169–226 (2004). DOI 10.1017/S0140525X04000056