Tracking common information and public announcements in online discussions. A data-driven and logically informed study of the Polymath-projects.

The problem that motivates this paper is the following: Given a data-set with records of interactions from collaborative science online, which background-theory should be adopted to study these digital traces if one's goal is to explain whether and how the collaboration was epistemically successful. I will approach this question on the basis of a specific case-study, namely the Polymath-projects initiated in 2009 by Cambridge mathematician and Field Medalist Timothy Gowers (see e.g. Allo et al. 2013). These are collaborative projects dedicated to specific research-level mathematical questions (finding a proof for a certain result). The centre of activity of these collaborations are interactions in discussion-threads on various weblogs, and the discussions in question are in principle open to anyone.

Building on the importance the logics for team-work developed in Dunin-Keplicz and Verbrugge (2010) accord to common knowledge and common belief, for instance by identifying common-belief as a necessary condition for common intentions and social commitments, I will argue that:

1. Although the data from the Polymath-projects suggest that methods from social-network analysis can be applied because digital traces of direct interaction and traces of temporal *co-presence* can be used to build interaction-networks and affiliation-networks, the direct application of commonly used methods and metrics in social network-analysis does not immediately allow us to detect how common belief can be achieved, which is exactly what we need to keep track of if we want to understand interactions as part of a collaborative enterprise.

2. The currently available logical methods for the analysis of social-informational phenomena are primarily concerned with information-diffusion caused by one-to-one interactions between connected individuals within social networks (Christoff and Hansen 2015, Liu et al. 2014). Furthermore, the formal analysis of the role of higher-order and common beliefs in social-informational phenomena is more often concerned with rational failures, like pluralistic ignorance (Hendricks 2010), than with epistemic successes. As such, the question of how common beliefs could arise when informational interactions are mediated by a social network has hitherto remained in the background.

In view of this, although insights from epistemic logic and logics for multi-agent systems are clearly relevant to the question of how we should understand successful collaborative science, it is still an open question of how these insights can be used when the available data initially only allow us to extract network-structures. An additional problem is that, because online discussions, as the ones found in discussion-threads on fora and blogs, are asynchronous, and effectively lead to uncertainty about who will read a given message, the public announcement required to obtain common beliefs, a necessary condition for successful team-work, seem impossible to achieve.

Because common belief is in practice often unattainable, even in situations where successful teamwork is required and attainable, Dunin-Keplicz and Verbrugge (2010) introduce the idea of an *awareness-dial* and argue that in practice only weaker forms of common belief are necessary for social commitments. This suggests that it may be sufficient to track approximations of common belief, and likely instances of public announcements in the available network-data. This can be done by identifying contextually defined sub-groups within which public announcements can be approximated.

A first way to do this is based on the notion of the “centre of discussion”. If we define the distance of a participant from the centre of discussion as proportional to the time elapsed since his/her last message, and assume that closeness of a participant to the centre of discussion correlates with how likely that participant will read a new message, then the open character of all messages (and meta-data like time-stamps) is sufficient to allow any participant to come to know who is likely to read a new message. If we assume that this is also commonly known, then the group of agents that are close to the centre of discussion can be said to share an informational context, and are hence able to make public announcements within this group.

A second way to do this is based on the identification of episodes or bursts of activity in discussions. By identifying such episodes and treating the participants to a single episode as a group within which during that episode public announcements can be made, we can identify probable public announcement without making assumptions about how close a participant has to be to the centre of discussion to read new messages or comments.

Each of these approaches is imperfect. Identifying episodes can for instance only take place after the fact (e.g. using a clustering algorithm), and can therefore not reflect the actual abilities of the participants. The centre of discussion metaphor, on the other hand, can only be used by deciding what it means to be sufficiently close to the centre. Yet, from a modeller’s perspective and given the available data, each can be put into practice and turned into an effective method that allows us to understand how a common understanding of the current state of a problem-solving process can be achieved, and how active participants can reliably keep track of how this common understanding evolves over time. Using data from the Polymath-projects it can be demonstrated how this can be achieved in practice, and which types of patterns this reveals.

While these proposals are useful in the context of the study of the Polymath-projects (an important case-study in the philosophy of mathematical practices), they are also fruitful illustrations of how higher-order and group-level information can be introduced in the study of social networks, and provide an alternative perspective on the question of how formal models of scientific communities should be empirically grounded (Martini and Pinto 2016).

References

Allo, P., Bendegem, J. P. V. and Kerkhove, B. V. (2013), *Mathematical Arguments and Distributed Knowledge*, Springer Netherlands, Dordrecht, pp. 339–360.

Christoff, Z. and Hansen, J. U. (2015), ‘A logic for diffusion in social networks’, *Journal of Applied Logic* **13**(1), 48–77.

Dunin-Keplicz, B. and Verbrugge, R. (2010), *Teamwork in multi-agent systems. A formal approach*, Wiley Series in Agent Technology, Wiley, Chichester.

Hendricks, V. F. (2010), ‘Knowledge transmissibility and pluralistic ignorance: A first stab’, *Metaphilosophy* **41**(3), 279–291.

Liu, F., Seligman, J. and Girard, P. (2014), ‘Logical dynamics of belief change in the community’, *Synthese* **191**(11), 2403–2431.

Martini, C. and Pinto, M. F. (2016), ‘Modeling the social organization of science’, *European Journal for Philosophy of Science* pp. 1–18.