Argument Mining in Social Media

A dominant characteristic of social media, just like in the wider web, is communication. Social media users are communicating; they are speaking and responding and sharing. However, identifying, extracting, and reusing information from otherwise unstructured inter-personal communication data is a task that involves solving many of the difficult problems that characterise computational processing of natural language. The sheer scale of data coming from social media compounds this issue with additional problems of how to deal with this so-called 'big data'. Perhaps natural language and big data techniques can be supplemented with approaches from other domains, such as that of argumentation theory [van Eemeren, 1995]?

Very often social media communication is competitive, if not outright antagonistic, but, ignoring for the moment explicit trolling, even vehement online disagreement can be useful, interesting, and valuable. This is because it often captures the real opinions of real people on the subject of important real-world topics. People often argue about things that are important to them. If these topics are important then there are usually groups, for example political, social and economic organisations, that are willing to take notice of the details of those arguments. However, at the scale of social media it becomes necessary to adequately structure these interactions so that they are readily reusable.

When we talk about communication, especially communication that involves a focus on disagreement or conflicting positions, then we are in the realm of argumentation rather than mere discourse analysis. By studying and modelling the arguments that people make online when using social media tools, and the ways that people interact to construct those arguments, we can build and deploy new tools that make it easier to extract useful information from social media data.

Therefore, perhaps instead of approaching social media mining from the perspective of either a big-data problem, or a natural language understanding problem, it can be approached as an argumentation theory problem; providing a domain in which new techniques can be deployed for identifying, eliciting, extracting, and structuring arguments. Furthermore, interaction techniques and user interface design can be used to make this process simpler with the aim of supporting the production of explicitly structured social media data in the form of arguments.

By taking this approach a number of advantages can be gained; In the short-term social media data becomes more immediately useful, especially from a

computational perspective. The content of explicitly structured arguments is more easily extracted, for example, to more easily distinguish not just sentiment, but also position and justification. Explicitly structured argument can also be more easily reused. For example in [Atkinson, 2012] arguments are stored using argumentation schemes [Walton,], stereotypical patterns of reasoning, which are used not just to populate the web interface but also both to manage and constrain the interactions available to users. An alternative to Atkinson's approach can be found in [Reed, 2007] where interaction is managed and constrained not by the types of arguments allowed by the available argumentation schemes, but by the kinds of things that can be said according to a governing protocol [Wells, 2012]. The specific type of protocol is a form of dialogue game [Hamblin, 1970], a multi-player, turn-taking game in which the players use their turn to speak, and thus enter content, in the form of arguments, into the dialogue. The range of things that can be said are constrained both by the protocol and the history of the dialogue as it develops. In [Reed, 2007] arguments are used to populate the knowledge-bases of intelligent agents which then interact with human-users via structured dialogue protocols. Arguments elicited via these protocols are subsequently reused as the knowledge-bases of newly instantiated intelligent agents such that each human user who interacts with the system extends the capabilities of the system and the resources available to subsequent users.

By taking a simplified initial approach, and by developing and exploiting new user interaction and interface techniques, we can subsequently provide the basis for more powerful automated techniques. In the longer term, the approach outlined above will provide large-scale corpora of analysed and explicitly structured argumentative data, and hence provide the foundation for new, automated, machine-learning approaches to mining argument structure from natural language.

[**Atkinson, 2012**] Wyner, A., Wardeh, M., Bench-Capon, T. and Atkinson, K.: A model-based critique tool for policy deliberation in Frontiers in Artificial Intelligence and Applications, Vol. 250. IOS Press (2012).
[**Hamblin, 1970**] Hamblin, C. L.: Fallaceies, Methuen and Co. Ltd (1970).
[**Reed, 2007**] Reed, C. and Wells, S.: Using Dialogical Argument as an Interface to Complex Debates. In IEEE Intelligent Systems Journal: Special Issue on Argumentation Technology 22 (6), pages 60-65 (2007).
[**van Eemeren, 1995**] Eemeren, F. H. van and Grootendorst, R. and Snoeck Henkemans, F.: Fundamentals Of Argumentation Theory, Lawrence Erlbaum Associates (1995).
[**Walton, 2008**] Walton, D., Reed, C., Macagno, F.: Argumentation Schemes. Cambridge University Press (2008).
[**Wells, 2012**] Wells, S. and Reed, C.: A Domain Specific Language for Describing Diverse Systems of Dialogue. In Journal of Applied Logic 10 (4), pages 309-329 (2012).