Microdebates for policy-making[*]

Simone Gabbriellini and Paolo Torroni

DEIS - University of Bologna
V.le Risorgimento, 2, 40136, Bologna - Italy
{`simone.gabbriellini;paolo.torroni`}`@unibo.it`

Web 2.0 platforms have rapidly become a mass phenomenon whereby billions of individuals consume and share resources. Scholars, advertisers and political activists see massive online social networks as a representation of social interactions that can be used to study or exploit the propagation of ideas, social bond dynamics and viral marketing, among others. Driven by a vision of a sustainable future and a more participated democracy, also administrations and policy-makers are more and more interested in using the Internet [1], and in particular the social Web, as an e-participation tool to support for public engagement and decision-making by the public, and for collecting, modeling, and presenting opinions and relevant information via usable interfaces [2]. Traditional surveys are a well developed and reliable method for retrieving opinions, but the selection of questions and the way these questions are posed may represent a crucial bias in faithfully representing citizens' opinions, in spite of the high standards for survey research that have been developed so far [3].

Web 2.0 platforms may contribute in mitigating such a bias, by allowing for online debates between informed citizens. The most famous virtual places to share opinions, e.g., FaceBook and Twitter, have a ludic and informal approach, however some others, such as DebateGraph, DBee, Debate.org, and Deliberatorium, have been developed with the specific aim to facilitate deliberation processes, by better organizing users' participation and decision-making processes in e-participation context.[1] Many other platforms are being developed within EU-funded projects, such as ePOLICY, IMPACT, and WEGOV.[2]

Thus online debates may be open to policy-makers's active contribution or simply to their inspection, and they could be used by the policy-maker as a powerful integration to questionnaires and polls. This approach to opinion gathering has the potential to overcome the above limitation of polls. In particular, citizens are free to exchange unconstrained opinions and critiques about certain topics. They can come up with new ideas and perspectives, as opposed to expressing preferences upon some predetermined options, and all in a bottom-up fashion [4]. However, the "freedom of expression" provided by online debates comes at a cost. In particular, it becomes very expensive for policy-makers to make sense of opinions emerging from comments contributed to online debates. In particular, he automatic extraction of opinions from text is all but a trivial task.

[1] `facebook.com`, `twitter.com`, `debategraph.org`, `dbeelife.com`, `debate.org`, `cci.mit.edu/research/deliberatorium.html`

[2] `epolicy-project.eu`, `policy-impact.eu`, `wegov-project.eu`

State of the art opinion mining/sentiment analysis techniques and tools look at sentiment orientation of opinions in terms of values in a positive/negative scale, typically by looking at corpora that include a certain number of sentences (e.g., online reviews about some product) [5][6]. Such an approach can provide very interesting results, especially if the domain is very well defined (e.g., a product, or a service). In fact, the classification accuracy in political debates is not yet as good [7] as that in other domains such as customer reviews [8], and, above all, it does not explicitly tell why certain opinions are in place and how they relate to other opinions.

Our work goes in the perspective of encouraging free, unconstrained online debate, as a tool in the hands of the citizens, who can use it to voice their opinions, and convey them to the policy-makers. To achieve this vision, we need to provide the policy-makers with tools to automatically make sense of possibly very lengthy online debates. Such tools should not only show the general sentiment around a specific topic, which is the approach of current sentiment analysis tools. Instead, they should also be able to **identify specific opinions, and the relations among them**. Such relations could be positive (support) or negative (counter). We identify **computational argumentation**, and in particular **abstract argumentation** [9], as the conceptual and computational framework to model opinions and reason from them automatically.

Current online debating tools, such as those we cited above, build on and extend the traditional forum-like structure, where users can reply or quote other users, by introducing debate-oriented concepts. They are not very different from a standard discussion forum with reputation, moderators and recommendation features. Moreover, they require the user to comply and adapt to the abstractions they are built around, and not vice-versa.

On the contrary, mainstream Web 2.0 social networking environments, e.g., Twitter, are very successful in achieving user engagement, by blurring the boundaries between ludic and serious [10]. **Our proposal is thus to develop an application based on a Twitter dialect that allows users to discuss about topics, aided** (in the back-end) **by computational argumentation**.

People use Twitter to talk about their daily activities and to seek or share information [11] by broadcasting brief textual messages (*tweets*) to people who "follow" their activity [12], in a micro-blogging fashion. Micro-blogging is a new form of communication whereby users can describe their current status in short posts distributed by instant messages, mobile phones, email or the Web [13]. We therefore introduce the concept of **microdebates**.

Microdebates are inspired by Twitter's microblogging character. Microdebates are sets of elements of content (e.g., *tweets*), each containing a contribution to a debate, such as an opinion, and may contain explicit references to other elements of content. Each element of content in a microdebate may use some combinations of characters (similar to hashtags) expressing positive or negative relations with other content elements. In this way, all that is asked of the user is to use certain combinations of characters in order to put their opinion in the context of other opinions. In exchange, users will have a way to convey

their perspective to the policy-maker: microdebates can be processed by automatic reasoners, such as argumentation-based reasoning tools [14]. The output of such tools may be visualized graphically, as clusters of coherent opinions around some topics, where different cluster may attack each other. Moreover, such output could be fed back to the user, in the same way as poll results are visualized (e.g., to those contributing to the poll). This could foster awareness of different opinions and encourage participation in a democratic policy-making process.

Twitter posts contain terms called hashtags, i.e. a # symbol followed by a text string, representing the stream of news the tweet belongs to. There may be more than one hashtag per post (in case the same post is related to multiple streams). Users on Twitter started the phenomenon of adding tags to their messages sometime around February 2008 [15]. This tagging behavior is distinct from those in other Web 2.0 systems, because users are less likely to index messages for later retrieval [16]. Tagging patterns in Twitter have a conversational, rather than organizational, nature [17].

In line with Twitter users' tagging behavior, we propose a Twitter dialect consisting of a custom set of tags to be used to annotate tweets in microdebates:

- a *hashtag* that will identify the discussion (e.g., #*debateName*): as customary, this ensures that the tweet will appear in the right stream (debate);
- one or more *annotations* using the $/!$ tags, where
  - $*opinionName* specifies the opinion this tweet supports, while
  - !$*opinionName* specifies the opinion this tweet counters.

The syntax for a microdebate is thus as follows:

⟨ microdebate ⟩ ::= ⟨ content element ⟩$^+$
⟨ content element ⟩ ::= ⟨ hashtag ⟩ ⟨ debate item ⟩$^+$
⟨ debate item ⟩ ::= ⟨ free text comment ⟩ | $⟨ opinionName ⟩ | !$⟨ opinionName ⟩

Notation ⟨ ... ⟩$^+$ indicates that the element in angle brackets is optional and that multiple occurrences are allowed. A *free text comment* is any free text not containing the special characters #/$. An *opinionName* is a tag given to a certain opinion; it should be formatted according to Twitter's tag syntax (alphanumeric strings with no spaces). The order of content elements in a debate, and of debate item inside a content element, is immaterial. This is how microdebates work:

1. content elements are *tweets* with a suitable *hashtag*, used to identify the micro-debates the users is contributing to. (Twitter then displays such tweet in the public stream associated with such a hashtag);
2. users annotate their tweets using $/!$ tags. When a user A specifies $*opinion$_1$*, it means that his comment supports *opinion$_1$*, which can be an opinion expressed by the user himself in the comment, or by another users B. In that case *opinion$_1$* will be seen as based on two comments, A's and B's respectively. The opinion name needn't be a summary of the user's opinion;
3. users can attack (counter) opinions using the !$ tag, e.g., by adding the !$*opinion$_2$* item in his tweet. This negation states that the tweet is a comment, which supports a certain opinion, and at the same time attacks *opinion$_2$*. This enables establishing relations amongst opinions;

4. if a user adds a tweet with a new $ tag, the user is in fact introducing a new opinion in the microdebate;
5. reply and re-tweets are handled like new tweets, thus personal replies are irrelevant to the debate.

We implemented a prototype using NetLogo [18] to extract relevant elements of content pertaining to a given microdebate from Twitter, and to compute clusters of coherent opinions (i.e., the *models*, in computational argumentation terms, defined based on specific semantics) based on such elements of content. We used Twitter API to retrieve tweets from Twitter, and Netlogo API to bundle our system into an extension with a basic parser that enables NetLogo to visualize and analyze the resulting argumentation framework.

References

1. Macintosh, A.: Characterizing e-participation in policy-making. Proc. HICSS, IEEE (2004)
2. Millard, J. et al. European eParticipation summary report. Tech. Rep., EC (2009)
3. Lynn, P.: Developing quality standards for cross-national survey research: five approaches. Int. J. Soc. Res. Method. **6**(4) (2003) 323–336
4. Toni, F., Torroni, P.: Bottom-up argumentation. Proc. TAFA, LNCS 7132, Springer (2012) 249–262
5. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retr. **2**(1-2) 1–135
6. Liu, B.: Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
7. Yu, B., Kaufmann, S., Diermeier, D.: Exploring the characteristics of opinion expressions for political opinion classification. Proc. D.Go (2008) 82–91
8. Kim, S.M., Hovy, E.: Automatic identification of pro and con reasons in online reviews. Proc. COLING/ACL 2006 483–490
9. Dung, P.M.: On the acceptability of arguments and its fundamental role in NMR, LP and n-person games. Artif. Intell. **77**(2) (1995) 321 – 357
10. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? Proc. WWW, ACM (2010) 591–600
11. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. Proc. WebKDD/SNA-KDD, ACM (2007) 56–65
12. Huberman, B.A., Romero, D.M., Wu, F.: Social networks that matter: Twitter under the microscope. First Monday **14**(1) (2009)
13. DeVoe, K.M.: Bursts of information: Microblogging. The Reference Librarian **50**(2) (2009) 212–214
14. Egly, U., Gaggl, S., Woltran, S.: ASPARTIX: Implementing argumentation frameworks using ASP. Proc. ICLP, LNCS 5366. Springer (2008) 734–738
15. Bruns, A., Burgess, J.E.: The use of twitter hashtags in the formation of ad hoc publics. Proc. ECPR, University of Iceland, Reykjavik (2011)
16. Romero, D.M., Meeder, B., Kleinberg, J.: Differences in the mechanics of information diffusion across topics. Proc. WWW, ACM (2011) 695–704
17. Huang, J., Thornton, K.M., Efthimiadis, E.N.: Conversational tagging in twitter. Proc. HT, ACM (2010) 173–178
18. Wilensky, U.: NetLogo. `http://ccl.northwestern.edu/netlogo/` (1999)