Large scale agreements via microdebates

Simone Gabbriellini and Paolo Torroni

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

Abstract. Argumentative debates are a powerful tool for reaching agreements in open environments. However, in large scale settings, such as social networks and massive multi-agent systems, making sense of ongoing debates may be a compelling task, and debates risk to lose their effectiveness. We thus propose "microdebates" to help organizing and confronting users' opinions in an automated way.

Keywords: abstract argumentation, negotiation in online debate, social networks, agreement facilitation.

1 Introduction

In the last decade, Web 2.0 platforms have rapidly become a mass phenomenon whereby billions of individuals consume and share resources. In such a setting, people became accustomed to arguing online in long-lasting debates, mainly in the form of comments in social network platform, such as FaceBook[1] and Twitter,[2] but also in the form of structured debates in *debate-friendly* tools. Among the latter we mention DebateGraph,[3] a powerful visualization tool; DBee,[4] a global debating network which features scoring and ranking with both positive or negative values; Debate.org,[5] a social network platform where users can start debate and comment with pro/cons rating against the main argument in the debate; and Deliberatorium,[6] a community-moderated system where comments need a moderator approval to be certified and visible by a larger community of commenters.

Indeed, Mercier and Sperber's argumentative theory of reasoning [1] tells us that people are good at reasoning when they communicate through an argumentative context. Arguments are used by communicants to convince other communicants, especially in absence of trust. When debating about an issues in these online settings, we thus expect that users will not only publish their

[1] http://www.facebook.com
[2] http://www.twitter.com
[3] http://www.debategraph.org
[4] http://dbeelife.com
[5] http://www.debate.org
[6] http://cci.mit.edu/research/deliberatorium.html

opinion (like in a review setting), but also try to convince others by producing arguments and rebut (attack) each others arguments.

Thus, argumentative debate seems to be a very promising tool for reaching agreement, with particularly interesting applications in a number of settings, including e-participation an policy-making. Indeed, a number of different platforms are being developed within EU-funded projects such as ePolicy,[7] whose aims include deriving social impacts through opinion mining on e-participation data extracted from the web; IMPACT,[8] which is developing an innovative argumentation toolbox for supporting open, inclusive and transparent deliberations about public policy; and WEGOV,[9] aiming to provide a toolset for exploiting existing social networking sites to engage citizens in two-way dialogs as part of governance and policy-making processes. The idea is that Web 2.0 platforms may overcome the limitations of traditional opinion gathering methods such as questionnaires and polls, by allowing for online debates between informed citizens, who can come up with new ideas and perspectives, as opposed to expressing preferences upon some predetermined options, and all in a bottom-up fashion [2]. However, the "freedom of expression" provided by online debates comes at a cost.

In particular, when we think of settings involving multitudes of interacting parties, such as social networks or large scale multi-agent systems, it becomes very expensive for by-standers and external observers to make sense of opinions emerging from online debates. An alternative approach could be to restrict oneself to getting a feeling of the general sentiment of an ongoing discussion, without necessarily having to really understand what is being said an why individuals make such and such claim and express such and such opinion.

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) [3][4]. Such an approach can be very effective especially if the domain is well defined (e.g., a product, or a service). In domains such as customer reviews [5] where the concepts involved can be defined in terms of specialized ontologies, and the jargon is pretty well defined and narrow, the classification accuracy of existing sentiment analysis algorithms is quite high. However, this is not the case in other domains, such as political debate [6]. Importantly, sentiment analysis 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. In said policy-making context, this could be 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

[7] `http://www.epolicy-project.eu`
[8] `http://www.policy-impact.eu`
[9] `http://www.wegov-project.eu`

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 [7], as the conceptual and computational framework to model opinions and reason from them automatically.

In computational abstract argumentation, as defined by Dung [7], an argumentation framework is defined as a pair $\langle X, A \rangle$, where X is a set of atomic arguments and A is a binary attacks relation over arguments, $A \subseteq X \times X$, with $\langle x, y \rangle \in A$ interpreted as "argument x attacks argument y." Collections of "justified" arguments can be described by various extension-based semantics [8].

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, such as Twitter, are very successful in achieving user engagement, by blurring the boundaries between ludic and serious [9]. 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 [10] by broadcasting brief textual messages (*tweets*) to people who "follow" their activity [11], 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 [12]. We therefore introduce the concept of microdebates.

2 Microdebates

Microdebates are inspired by Twitter's microblogging character. A microdebate is a stream of tweets where users annotate their messages by using some special tags. 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 [13]. Twitter tagging behavior is distinct from those in other Web 2.0 systems, because users are less likely to index messages for later retrieval [14], and this is reflected by the fact that tagging patterns in Twitter have a conversational, rather than organizational, nature [15].

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 (microdebate);
– one or more *annotation(s)* 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:

$$
\begin{aligned}
\langle \text{ microdebate } \rangle &::= \langle \text{ content element } \rangle^{+} \\
\langle \text{ content element } \rangle &::= \langle \text{ hashtag } \rangle \ \langle \text{ debate item } \rangle^{+} \\
\langle \text{ debate item } \rangle &::= \langle \text{ free text comment } \rangle \\
&\quad \mid \$\langle \text{ opinionName } \rangle \\
&\quad \mid !\$\langle \text{ opinionName } \rangle
\end{aligned}
$$

Notation $\langle \ldots \rangle^{+}$ indicates that multiple occurrences of the element in angle brackets 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. An example of a hypothetical microdebate is shown in Figure 1.

A microdebate is thus a set of elements of content (such as *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 receive a help in making sense of a (possibly lengthy) debate: microdebates can be processed by automatic reasoners, such as argumentation-based reasoning tools [16] and the output can be visualized graphically as clusters of coherent opinions, where different cluster may attack each other. This could foster awareness of different opinions on a topic and encourage arguers to reach an agreement.

This is how microdebates work:

1. content elements are *tweets* with a suitable *hashtag*, used to identify the microdebates users are 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 is abstract, and does not need to be a summary of the user's opinion;

Fig. 1. An example of microdebate on Twitter. $ tags and !$ tags represent arguments and attacks between them.

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 (unless they contain $/!$ tags that are meaningful for this debate).

3 Microdebates at work

We implemented a first prototype of the system as an agent-based model in NetLogo [17]. In this model, each agent represents an argument used in the microdebate. Attacks between arguments are represented by directed links from an agent to another one. We used the Twitter API to retrieve tweets from Twitter, and the Netlogo API to bundle our system into an extension with a basic parser (called *microdebate*), that enables NetLogo to visualize and analyze the resulting argumentation framework.

As a first step, we extract and parse the stream of tweets in a selected microdebate, so that we have:

- for each new $$opinionName$ tag, a new argument is created;
- for each new !$$opinionName$ tag, a new attack link is created against the named argument

To retrieve the microdebate, it suffices to enter a debate identifier, in the form $$debateName$, in the GUI's *debate* text box (see Figure 2). In our example, the debate identifier is $$energyalt$. Of course, there is a difference between an opinion and an argument, the former being a claim without evidence, the latter being a claim with evidence (supported to convince others that the claim is supported). At the same time, not all the comments expressed by users can turn out to be "well-formed" arguments. Nevertheless, at this stage, we turn every $$opinionName$ into an argument belonging to a preliminary argumentative framework that we define *naive*.

In order to improve our framework in this respect, we store inside each argument all the free text comments that refers to $$opinionName$ in the microdebate. We then propose argument classification as a way to verify if each claim is a well-formed argument or not (see Figure 3):

- if, based on the comments it contains, the claim proves to be indeed a well-formed argument, we keep it in the argumentative framework;
- otherwise, if based in its comments, the claim proves not to be a well-formed argument, we exclude it from the argumentative framework.

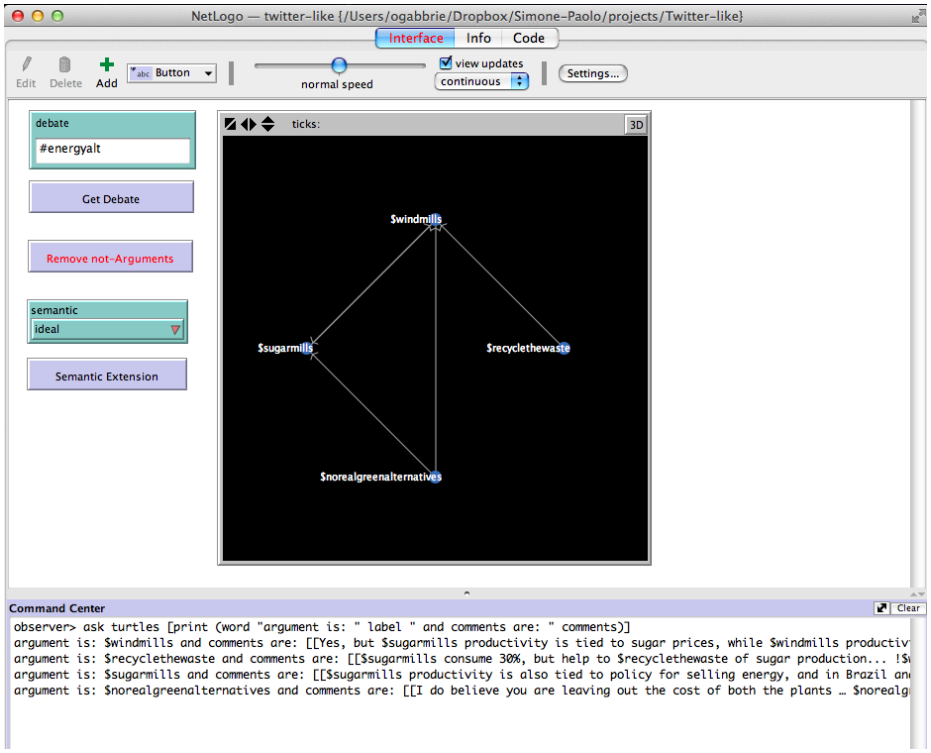

Fig. 2. From a Twitter microdebate to an argumentation framework.

This method allows to obtain a polished up argumentation framework, where all (and only) well-formed arguments are retained. Being this an initial prototype, such processing is currently made by hand. In the concluding section of this paper we elaborate on how we plan to improve this stage.

Once we have only arguments and attacks among arguments, we can compute semantic extensions on the argumentative framework.

Our prototype can compute extensions based on a variety of semantics, including admissible, complete, grounded, ideal, preferred, stable, semistable, and stage semantics.

In Figure 4 the *complete* semantic extension has been calculated, that states: a set of arguments is a semantic extension iff the set include all the arguments that it defends. As we can see, the two arguments $\$sugarmills$ and $\$recyclethewaste$ are the winners over $\$windmills$, demonstrating that the stream of tweets that compose the microdebate $\#energyalt$ can be summarized in a very compact and efficient way.

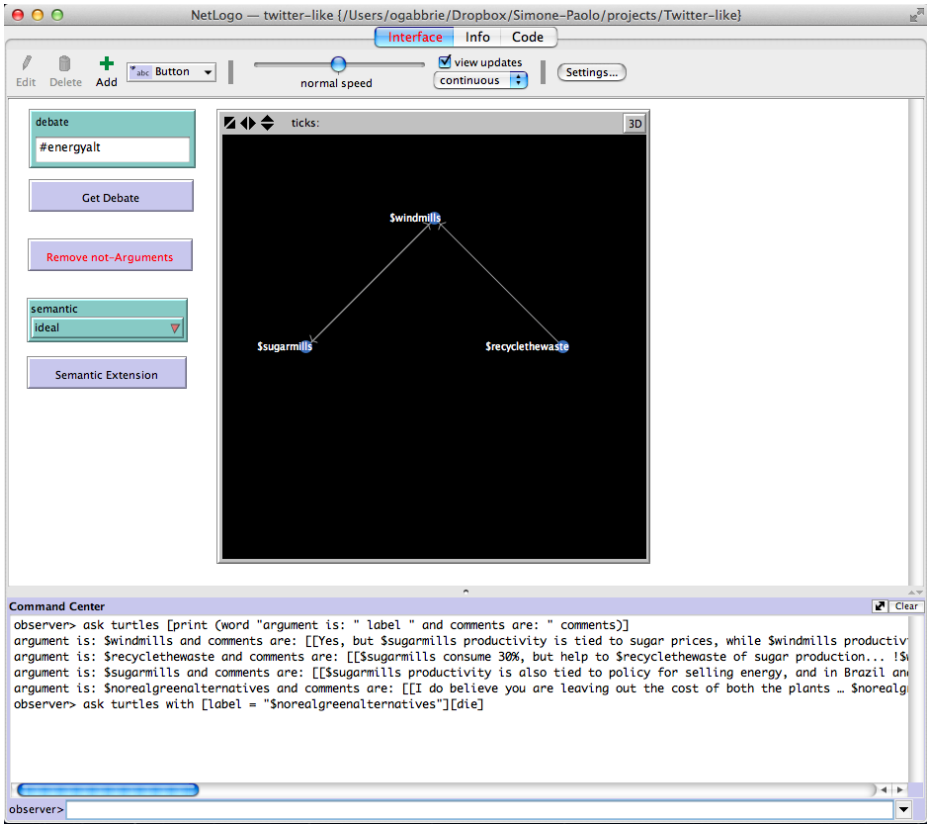

Fig. 3. Polished up argumentative framework

4 Conclusions

The purpose of our proposal is to help reaching an agreement in a debate by formalizing and rationalizing a debate. Recent findings in cognitive science [1] suggest that people are good at arguing, actually that the main function of reasoning is argumentative. However, when big numbers are in play, it may be difficult for by-standers and potential contributors to make sense of online discussions. By microdebates, we aim to help people understand how a topic is being discussed, what positions (arguments) are involved in the debate, and what are the relations of attacks between such arguments. Ultimately, we aim to help people argue in a better way, defend their reasons and learn how to rebut each other's attacks.

The "microdebates" we propose follow the bottom-up argumentation philosophy introduced in [2]. Users contribute to a debate by sending out annotated comments, and as a result, arguments arise bottom-up. In particular, it is not

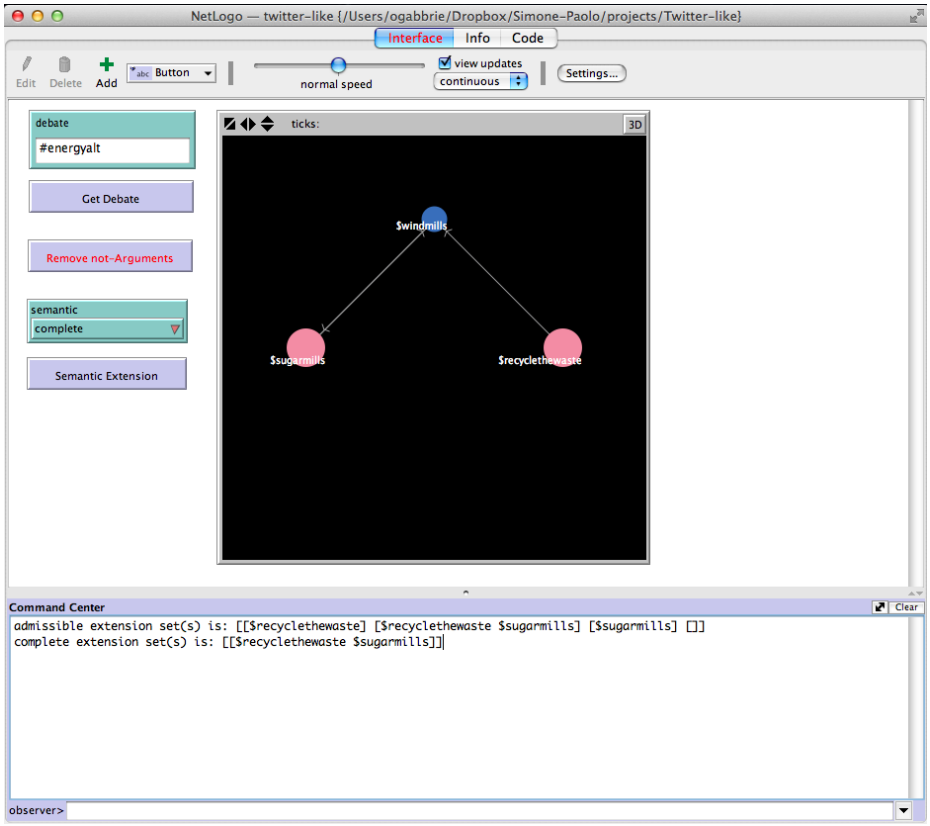

Fig. 4. Visualisation of the complete semantics

necessary that the same user defines a well-formed argument, because many users can contribute, tweet by tweet, to support the same argument, by adding elements that turn a claim into a well-formed argument, or by finding rebuttals and counter-attacks to the arguments in place.

All this effort should help to produce a more relevant discussion, in a very user-friendly way: users can annotate their messages in an everyday-life style, and they do not have to conform to (rigid) rules of another *debate-oriented* interface.

The computational underpinning of our proposal is abstract argumentation, and is orthogonal to the choice of an extension-based semantics, by design. Different semantics may suit to different applications in various ways.

Our tool is partially implemented. Two NetLogo extensions are already implemented: *microdebate*, for the processing of tweets, and and *arguments*, for computing extensions. We are still at an early stage in argument classification, whose purpose is to filter arguments and keep well-formed ones only. For that,

we plan to use third-party semantic tagging tools, such as COGITO.[10] Our idea is to define what a "well-formed argument" is by way of COGITO rules, and delegate to a COGITO module a fully automated argument filtering process.

We also plann to extensively test our method with case-studies, in order to understand the effectiveness of this approach in a real-world setting. In particular, we are designing tests with debates concerning renewable energy sources, environment and sustainability, in the context of the above-mentioned ePolicy EU project.

The research presented here is high-risk, because many innovations are required all together for this to succeed. For instance, using our syntax,Twitter users may develop habits that could be different from what we expect, leading to unforeseen system behavior. The parser will probably need to get adjusted once some data from a case-study has been retrieved.

Moreover, since our method (and bottom-up argumentation in general) needs active engagement from users, we could end up with poor data to analyze if our users will not get *truly* involved in the process. However, we hope that this factor may be mitigated by a unique feature of microdebates: they do not require a dedicated platform. Users do not need to learn and get accostumed to new interfaces, because microdebates are only based on tweets (as opposed to graphic items, such as bubbles and links).

Having said that, we believe that the strengths and potential of our approach overcome its limitations. First, microdebates allows deep analysis of arguers position in a debate, an important step toward the reaching of an agreement between arguers. Furthermore, by-standers may be encouraged to participate, since they have a clear visualization of what is happening in the debate - and what position (arguments set) is going to dominate or get defeated.

Second, there is no need to manually analyze documents, because posts are annotated by users. This form of *crowdsourcing* reduces the amount of qualified labor needed. An important bottle-neck is argument classification, but we hope to be able to set up automated procedures for it.

Third, the microdebate approach develops a technology that may be useful in many other domains, because it is based on a multi-disciplinary approach that well suits the needs of diverse domains where debates are allowed, such as policy-making, Moreover, such technology, initially developed for human-to-human interaction, may as well be exported to software agents. We can think of agents communicating with one another in a tweet-like fashion, and new algorithms could be developed to automatically reach agreements between agents on a variety of domain. This may open a promising strand of research.

Fourth, our approach exploits the so-called wisdom of the crowds (as in *bottom-up argumentation*): arguments arise bottom-up from the debate and it is not necessary for a single user to express the argument entirely, because other users can contribute to the same argument. Finally, it has an open approach that allows all users to visualize dynamically the outcome of the analysis.

[10] `http://www.expertsystem.net/products-technology/cogito-semantic-tagger`

Acknowledgments

We thank the anonymous reviewers for their useful and encouraging feedback.

This work was partially supported by the ePolicy EU project FP7-ICT-2011-7, grant agreement 288147. Possible inaccuracies of information are under the responsibility of the project team. The text reflects solely the views of its authors. The European Commission is not liable for any use that may be made of the information contained in this paper.

References

1. Mercier, H., Sperber, D.: Why do humans reason? arguments for an argumentative theory. Behavioral and Brain Sciences **34**(02) (2011) 57–74
2. Toni, F., Torroni, P.: Bottom-up argumentation. In Modgil, S., Oren, N., Toni, F., eds.: TAFA 2011: Revised Selected Papers from the First International Workshop on Theory and Applications of Formal Argumentation. Volume 7132 of Lecture Notes in Computer Science., Springer (2012) 249–262
3. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retr. **2**(1-2) (January 2008) 1–135
4. Liu, B.: Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
5. Kim, S.M., Hovy, E.: Automatic identification of pro and con reasons in online reviews. In: Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions. (July 2006) 483–490
6. Yu, B., Kaufmann, S., Diermeier, D.: Exploring the characteristics of opinion expressions for political opinion classification. In: Proceedings of the 2008 international conference on Digital government research. dg.o '08, Digital Government Society of North America (2008) 82–91
7. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2) (1995) 321 – 357
8. Baroni, P., Giacomin, M.: Semantics of abstract argument systems. In Simari, G., Rahwan, I., eds.: Argumentation in Artificial Intelligence, Springer-Verlag (2009)
9. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? In: WWW '10: Proceedings of the 19th international conference on World wide web, New York, NY, USA, ACM (2010) 591–600
10. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, New York, NY, USA, ACM (2007) 56–65
11. Huberman, B.A., Romero, D.M., Wu, F.: Social networks that matter: Twitter under the microscope. First Monday **14**(1) (January 2009)
12. DeVoe, K.M.: Bursts of information: Microblogging. The Reference Librarian **50**(2) (2009) 212–214
13. Bruns, A., Burgess, J.E.: The use of twitter hashtags in the formation of ad hoc publics. In: 6th European Consortium for Political Research General Conference, University of Iceland, Reykjavik (August 2011)

14. Romero, D.M., Meeder, B., Kleinberg, J.: Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In: Proceedings of the 20th international conference on World wide web. WWW '11, New York, NY, USA, ACM (2011) 695–704
15. Huang, J., Thornton, K.M., Efthimiadis, E.N.: Conversational tagging in twitter. In: Proceedings of the 21st ACM conference on Hypertext and hypermedia. HT '10, New York, NY, USA, ACM (2010) 173–178
16. Egly, U., Gaggl, S., Woltran, S.: ASPARTIX: Implementing argumentation frameworks using answer-set programming. In Garcia de la Banda, M., Pontelli, E., eds.: ICLP: Proceedings of the 24th International Conference on Logic Programming. Volume 5366 of Lecture Notes in Computer Science. Springer (2008) 734–738
17. Wilensky, U.: Netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. (1999)