Sketching the vision of Debate Web

Antonis Bikakis[1], Giorgos Flouris[2], Theodore Patkos[2], and Dimitris Plexousakis[2]

[1] Department of Information Studies, UCL, UK
a.bikakis@ucl.ac.uk
[2] Institute of Computer Science, FO.R.T.H., Greece
{flouris,patkos,dp}@ics.forth.gr

Abstract. Web users have changed the Web from a means for publishing and exchanging documents to a Web of Opinions. Current web technologies fail to support this change: arguments and opinions are uploaded in purely textual form; as a result they cannot be easily processed and utilised. This paper describes the vision of Debate Web, which will enable the extraction, discovery, retrieval, interrelation and visualisation of the vast variety of viewpoints that already exist online, based on machine-readable representations of arguments and opinions.

1 Motivation

People today use the Web to make comments, express their opinions, participate in debates and discuss subjects of any matter in blogs, wikis, forums and social networks. This information is published and stored in textual form; therefore it cannot be easily retrieved, processed and interlinked. The Semantic Web [1] and the recent linked data hype overcome some of the limitations of natural-language Web pages using appropriate methodologies for posting and interlinking semantical information in a machine-readable way. Their focus, though, is on the representation of data, rather than arguments or structured opinions. Computational argumentation [2] relies exclusively on the formal properties of logical arguments and factual information, and cannot satisfy the primal reason why opinions reach the Web in the first place, which is to be *persuasive*. This feature is important, in order to create *debates with a purpose*, where each argument is formulated with a certain aim: to persuade a certain audience on some topic.

We envision a new version of the Web, which we call *Debate Web*, which will depart from the processing of purely logical arguments and will move closer to human understanding. It will be based on machine-readable representations of arguments, opinions and any information that is related to their believability such as evidence, provenance, audience profiles, emotions, trust, context and popularity. It will use appropriate languages to describe this information; appropriate tools that will enable people to upload, retrieve and associate it; and appropriate technologies to extract, store, interlink, analyse and present it. The ultimate outcome will be to offer the means for assisting humans in participating in online debates and collective decision making processes in a persuasive manner.

2 Realising the Vision

There are two main fields that need to be explored or advanced for the realisation of this vision: *realistic argumentation* and *web technologies*.

Understanding Realistic Argumentation. Debate Web is a lot more than an argumentation system deployed in a global scale. The main challenge here is the shift from logical argumentation (the object of study for computational argumentation) to realistic argumentation. Realistic argumentation does not only appeal to the logic of the audience, but also to its emotions. It is only partly based on facts and data, often employing additional techniques such as the clever use of verbal cues and the semantic structure of text/speech (politeness, aggressiveness etc), as well as different argument schemes based on factors such as appeal to authority or expert opinion, popularity of supported claims, peer-pressure, analogous arguments, proof-by-example, non-logical (e.g., statistical) correlations between different arguments, and others [4].

Web Techologies. Debate Web will reuse technologies from the current Web such as the web infrastructure and protocols (e.g. TCP/IP); the data model and ontology languages of the Semantic Web (e.g. RDF/S, OWL); and techniques such as social tagging, crowdsourcing, voting and others from the Social Web and the Pragmatic Web. On top of these, it will provide higher-level (and therefore more useful) information services to its users as a result of the analysis, association, aggregation and summarisation of different types of semantically annotated data: from raw data to structured arguments and opinions. To achieve this the following challenges need to be addressed:

- **Extraction and Annotation.** Argument and Opinion mining and NLP techniques must be employed to enable the identification and extraction of arguments and opinions out of text; and techniques such as crowdsourcing and annotation tools will enable users to contribute to this process. and annotation.
- **Representation and Interchange.** The development of a semantically explicit representation model for arguments and opinions (e.g. in the form of an ontology) will allow different independently developed applications to comprehend them in a common manner and interoperate within an integrated environment.
- **Storage and Management.** Specially designed knowledge bases will be needed to store arguments, opinions and related metadata. They must also provide: query and inference support; support for updates, repairs and change monitoring; alignment with related ontologies; and propagation of information among different systems.
- **Reasoning and Analytics.** Appropriate analysis and reasoning mechanisms will enable users to retrieve information by posing structured queries, and perform sophisticated aggregation and summarisation operations. More complex forms of reasoning will allow the identification of implicit relationships between arguments, or the development of new forms of acceptability semantics along the tradition of [3].
- **Presentation and Visualisation.** Given the sheer size of the Web, a large number of arguments and opinions on most topics will be available. Presenting everything to the user is certainly not productive. Some kind of selection, aggregation or summarisation, which takes into account the user's information needs or background is necessary, along with a ranking process that selects the most persuasive ones.

Figure 1 displays the relevant research fields and technologies in a two-axis chart. The position of the technology along the horizontal axis represents both the current and the required maturity of each technology to solve the respective challenge. The left side of each rectangle represents the current capacity of the corresponding technology to address the related challenge, whereas its right side represents additional advances that need to be achieved (and how far in the future these are estimated to occur) before actually solving the respective challenge in its entirety. The vertical axis represents the kind of progress required per technology to overcome the respective challenge.

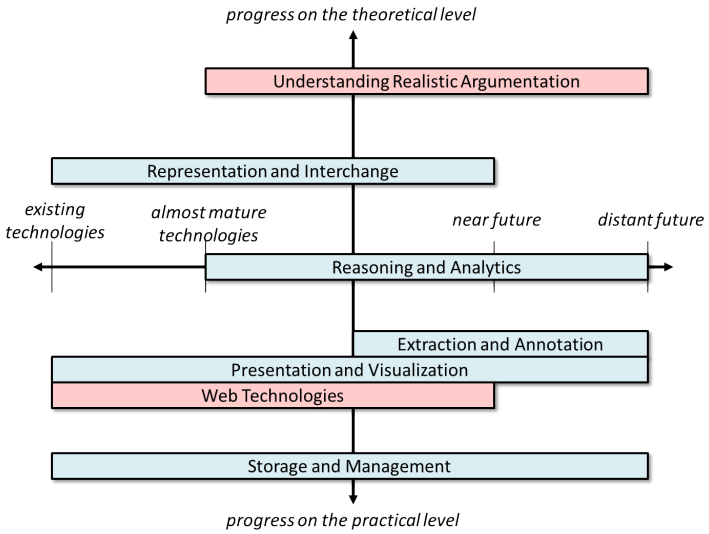

Fig. 1. A 2-dimensional categorisation of related technologies and challenges

Acknowledgements

We would like to thank Leon and Emil for the inspiring discussions we have had on some of these topics since Antonis joined the ICR team in September 2010, as part of a research proposal on persuasion.

References

1. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American (2001)
2. Besnard, P., Hunter, A.: Elements of Argumentation. The MIT Press (2008)
3. 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), 321–357 (1995)
4. Walton, D.: Fundamentals of Critical Argumentation. Cambridge University Press, Cambridge, UK (2006)