Open Research Online

The Open University's repository of research publications and other research outputs

Improving online deliberation with argument network visualization

Conference or Workshop Item

oro.open.ac.uk

Improving Online Deliberation with Argument Network Visualization

Anna De Liddo
KMi, Open University
Milton Keynes, UK
anna.deliddo@open.ac.uk

Simon Buckingham Shum
KMi, Open University
Milton Keynes, UK
simon.buckingham.shum@open.ac.uk

ABSTRACT

Social media are increasingly used to support online debate and facilitate citizens' engagement in policy and decision-making. Nevertheless the online dialogue spaces we see on the Web today typically provide flat listings of comments, or threads that can be viewed by 'subject' line. These are fundamentally *chronological* views which offer no insight into the *logical* structure of the ideas, such as the coherence or evidential basis of an argument. This hampers both quality of users' participation and effective assessment of the *state of the debate*. We report on an exploratory study in which we observed users interaction with a collective intelligence tool for online deliberation and compared network and threaded visualizations of arguments. We contend that animated argument networks enhance online debate reading when data complexity increases, improve understanding of the argumentation data model and promote users engagement by improving users emotional reactions to the online discussion tool.

Author Keywords

Collective Intelligence, Argumentation, Online Deliberation, Visualization, Evaluation

INTRODUCTION

One of the Web's most phenomenal impacts has been its capacity to connect and harness the ideas of many people seeking to tackle a problem. For instance, social media appear to have played specific and significant roles in helping communities form and mobilize, even to the level of political uprisings [18]. One of the emerging terms for this phenomenon is *Collective Intelligence* (CI), reflected in the first conference devoted to the topic in 2012 (ci2012.org) and active research in several fields [9]. However, while most common CI systems make it as simple as possible to begin contributing to a collective debate, they have serious limitations when it comes to aggregating analyses and arguments at scale. The online discourse spaces we see on the Web today typically provide for the addition of flat listings of comments, listed by date (e.g. comments in Facebook; on web articles; on blog posts), or threaded in a strict tree which can be additionally viewed by 'subject' line (e.g. Google or Yahoo Groups; Forum-Software.org). These are fundamentally *chronological* views of the unfolding conversation drawing attention to the most recent utterances, but offer no insight into the *logical* structure of

the ideas, such as the coherence or evidential basis of an argument. At a glance, all one can tell us which threads have most contributions, or have been most active. In essence, because very simple ontologies underpin these platforms, there is no way for a machine or person to *assess the state of the debate*, other than reading huge numbers of contributions. The challenge of designing usable, enjoyable platforms with richer representational schemes, which add real value to user experience and performance, is at the core of a 25 year research programme dating back to pre-Web hypertext systems [2].

Contested Collective Intelligence

In all but the most trivial dilemmas, rarely does everybody agree, for instance, due to different intellectual traditions, competing organizational objectives, information overload or ambiguity. Building common ground by understanding and overcoming differences is at the core of what it means to transform a set of isolated individuals into an effective collective. Consequently, the concept of *Contested Collective Intelligence (CCI)* has been recently defined as a distinctive subset of CI [5]. Combined human-machine annotation and analytics are needed for the many societal contexts in which deliberation and debate are the hallmarks.

Computer-Supported Argument Visualization (CSAV)

An established research literature documents the advantages, and challenges, of making the structure and status of a dialogue or debate more visible [2]. Following Concept Mapping [14], CSAV makes visually explicit users' lines of reasoning and (dis)agreements. Naturally, the use of semantic networks to provide computational intelligence has now converged with Web Science, resulting in *Web 2.0 Argumentation* [3] and a semantic web standards-based *Argument Web* [15]. Many CSAV tools are now available, exhibiting a wide mix of network visualizations and more conventional threaded renderings. *Deliberatorium* [11] uses a linear-threaded visualization, while *Cohere* [3] supports a graph visualization of the online discourse. *Debategraph* [6] and *CoPe_it!*[21] enable users to switch between a linear, threaded and graph views. As with most web applications, design decisions appear to have been made more on intuition and craft skill. Experience tells us that as the number of participants and nodes grows, and if more expressive ontologies are used to model the discourse, network

visualizations can become unwieldy. However, to date, we have not yet found a systematic evaluation of the merits or otherwise of CSAV vs. threaded interfaces. The internet usage pattern of "participation inequality" [10] confirms that typically, only a very small percentage of users contributes to a website, compared to vast majority who consume. However, that 1% of users are still critical, and one strand of CSAV research has focused on the learning curve for content creation with such tools, since they require more effort than simply pressing *Add Comment* or *Reply*. Our prior studies of graphical argumentation show there is a complex interplay between the expressiveness of the notation, user interface affordances, user expertise, task strategy, and synchronous vs. asynchronous mapping [4, 16, 17]. While browsing, searching, reading and writing clearly cannot be divorced from each other, the focus of the study reported here is not on the authoring process, but on the experience and performance of the majority of users who will be *reading and searching* the online deliberation platforms.

RESEARCH QUESTION AND METHODOLOGY

Does an interactive, self-organizing network visualization of arguments provide advantages over a more conventional threaded interface for reading and search? The following research questions was formulated:

RQ: What are the affordances of threaded and network visualization of arguments for supporting information seeking tasks?

A grounded theory analysis has been conducted to analyze the effects of using different graphical interfaces for the representation of arguments, and to assess to what extend these affect the way in which users read and understand the online discourse. Grounded Theory is a qualitative research method [8] which has been used to provide insights on people's views, behaviors, understanding and experience with technology in HCI research and very recently the discussion has been revamped on what are the nuances and innovation to the method that may help to respond to the increasingly request and diffusion of qualitative research studies in HCI [1].

USER STUDY

Goal, Design and Participants

We studied an heterogeneous group of researchers and practitioners in Higher Education engaged in the use of a collective intelligence tool for collaborative argumentative discourse and knowledge construction. An exploratory user study was run to observe users' performance under three information-seeking tasks, and compare their performances using two different user interfaces for arguments visualization (threaded vs network visualization of arguments). Participants were divided in two groups of 5. The 10 subjects were drawn from the members of different [anonymized] University departments, so with widely mixed IT expertise. They were randomly allocated to the two different groups, but we verified post-hoc that IT expert/non expert ratio in each group was approximately the same. The median age was 40 (with range from 32 to 48) with the majority of users being either native or near-native English speakers. Most subjects were employees at the [anonymized] University, all having university degrees or higher.

Online Environment: The Evidence Hub

The online discussion tool used throughout the study is the Evidence Hub system. The Evidence Hub is a collective intelligence and online deliberation tool to support argumentative knowledge construction by crowdsourcing contributions of issues, potential solutions, research claims and the related evidence in favor or against those. The Evidence Hub is particularly oriented to support community of practitioners to build evidence-based knowledge about specific key challenges that are set up to the community. This *key challenges* can then be addressed by tackling specific sub *issues*, that the community contribute to add to the system. Each sub issue can then be tackled by proposing solutions to it, or by sharing specific research claims that can help tackling the issue. Finally *potential solutions* and *research claims* can be debated by the community by advancing *evidence* in favor or against those and by providing relative *resources* backing them up. By scaffolding users contributions in this way, the Evidence Hub aim to effectively crowdsource and support large-scale deliberation in e-democracy and decision-making processes. To allow comparison between users interaction with two graphical representations of arguments, two different versions of the evidence Hub have been set up which used different user interfaces for arguments visualization. The two versions of the Evidence Hub though pointed at the same database to make sure that participants in the two groups would receive exactly the same quantity and type of information.

The Linear-threaded Interface for Arguments Visualization
The linear-threaded interface was used by Group1(Figure 1). This interface for arguments visualization is similar to the most common threaded online discussion interfaces to support argumentative discourse activities (se in example tools such as MIT Deliberatorium [11]). It consists of a classical threaded discussion visualization in which issues set the focus of discussion (title of the page in Figure 1), and then potential solution are listed below (light bulb icon are placed before each solution). Each solution can then be expanded on demand. Progressive indentation of text shows the supporting and challenging evidence for this solution, and the resources supporting them (see Figure 1).

The Network Interface for Arguments Visualization
The network interface was used by Group 2: This consists of an argument map built by following the modified IBIS,

PHI model of arguments [13]. Each statement added to the discussion is here represented as a node in a semantic graph structure (Figure 2). In this visualization, node icons and colors represent the rhetorical role of the statement in the discussion: issues are characterized by a red question mark icon in a dark turquoise node (yellow in Figure 2 because the node is selected); solutions are distinguished by a light bulb icon (see light green nodes), while pro and con are characterized by purple nodes and are connected with a green link or red link which respectively indicate "supporting" or "challenging" relations.
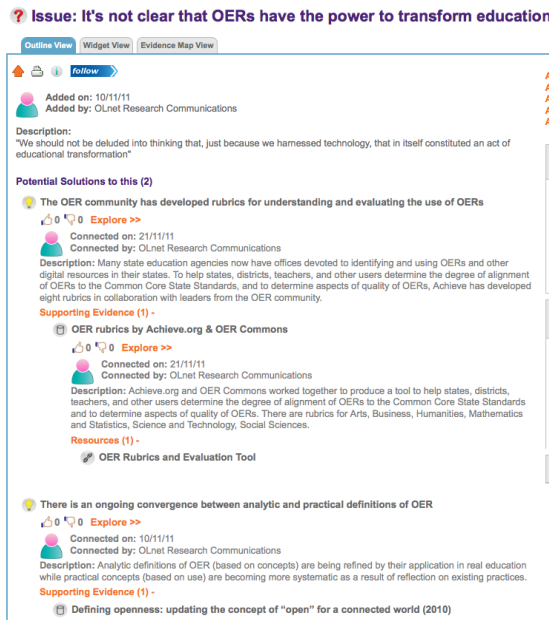

Figure 1. Screenshot of the linear-threaded Interface for Arguments Visualization

The network visualization of the Evidence Hub is build with a Java Applet, which lays out the nodes dynamically, and following a gravitational algorithm. This adds a specific animation component to most common argument network visualization tools.
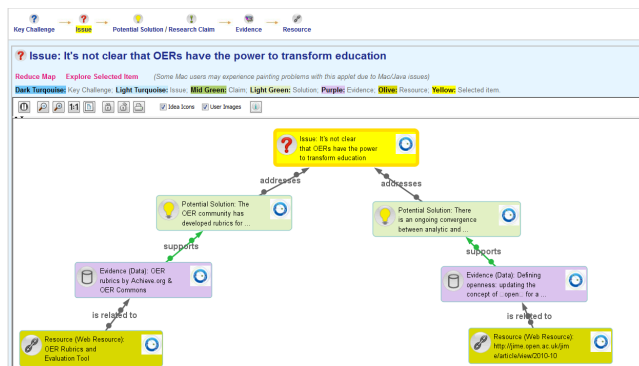

Figure 2. Screenshot of the Network Interface for Arguments Visualization

DATA ANALYSIS AND RESULTS

The experimentation design aimed to provide an answer to the RQ by focusing on three specific information-seeking tasks (for issues of space we skip detailed description of the tasks):

1. Identifying solutions to an issue (Task1)
2. Identifying synergies between solutions (Task 2)
3. Identifying contrasts in the wider debate (Task 3)

To compare user's performance across these three tasks we used three metrics: *Task Accomplishment*; *Data Model Interpretation* and *Emotional Reactions* to the two interfaces. In the following we analyze and compare users' performance by focusing on each of these metrics.

Task Accomplishment

By filtering clips by "performance" code, we found that there are 4 main sub-categories of performance emerging form the analysis, that are: *Accomplished easily*, *Accomplish but incorrect*, *Not accomplished*, and *Give Up the task*. From the analysis of the quotations and of the screen capture video of the interaction we found that there are two cases in which participants could not accomplish the task. The reason of task failure is neither to ascribe to the type of interface for graphical representation of arguments or to problems of data model interpretation. Participants did not accomplish the task because they did not understand the question posed by the moderator. See i.e. quotes from P9 and P10:*"..sorry I do not understand the question...." "What do you mean? I do not understand…".*

Moreover, two interaction events have been coded as "*accomplished but incorrect*". Both events occurred in Group1 (linear interface). Participants who experienced those events struggled to find the answer because this required them to click and expand two argumentation sequences and then read and compare them. This is not convenient with a linear interface because it requires a lot of clicking, expanding, scrolling up and down the screen, and a lot of reading. In a linear interface the structure does not help to interconnect and compare content and this is a burden to the user. Because of this reason some users focused more on the content rather than on the argumentation process, which lead to digressions and incorrect responses.

On the contrary the interaction events coded as "*Accomplished easily*" were experienced by 4 participants in Group 2 and 2 participants in Group 1. The reasons why Group 2 could accomplish easily are mainly due to the visual hints provided by the network representation. Specifically, links labels and colors seem particularly predominant in determining success. Users that accomplish the task easily usually relay on links colors and link label paying less attention to the iconography of the nodes. The number of links (that is to say connections density) seems also to be a very effective way to provide answer to the task. Interestingly visual hints were effectively used not only by participants that have self

declared to be very visual thinkers, but also by participants that have confessed they are "*not visual and prefer text to graphical representation*". See in Example the case of P8:

"*..here we are, this one...is that correct?.... I was not concentrating on what part of the diagrams are solutions and what constitutes evidence ... I was just looking for intersecting lines and I was thinking that have to be there where there are converging lines...so I thought may be is that one there and I zoomed to check...*"

Comparing Data Model Interpretation supported by the two interfaces

In order to compare linear and network visualization of arguments in term of how they supported Data Model Interpretation (DMI) we first identified the emerging codes related to DMI and merged them under the 5 main classes (*get the model, get the model after a while, get the categories right, get the model but is unsure, misinterpret categories, does not get the model*). We then analysed the interaction events and quotations for each main class.

Results show that DMI was better supported by network visualization of arguments. Category misinterpretation and uncertainty in data model interpretation tend to occur more frequently in Group1 than in Group2, while network visualization of arguments seems to support a complete and correct understanding of both categories and data model. Findings seems to suggest that one possible motivation for this is that argument maps provides examples of how the content of discussion is interpreted in the data model supported by the Evidence hub. In this way users can learn by example ("*let's see what other people have put under this category*" - P8) and at each exploration of a new argument map they reinforce their understanding of the data model ("*this is my solution and this evidence support the solution, and these two nodes challenges it...and the clue is the red "challenges" link* - P7)".

Comparing Emotional Reactions to the Two Interfaces

Emotional reaction to the graphical representation of arguments is another emerging category that we used to measure users satisfaction and emotional attitude toward the two proposed interfaces. Results show that a general sense of surprice and positivity toward the network visualization is recorded, which easily sparks into linkeness and even exhitment ("*ahhh so fabulous!*"-P7; "*wow this is amazing!*"-P6). This also seems to increase users confidence with the tool ("*I feel confident, I's pretty sure this is the answer*"-P7). The main objects of surprise and exhitement toward the network visualizations are usually due to the self-arranging graph applet which presents the map as floating nodes, slowly arranging on the screen("*it is like a gelly, it is so fantastic!*"-P7; "*it is all shifting! It is interesting... I quite like that!*"-P8). Movements provokes surprise and exhitement and it is also recognized as a useful hint. By looking at the floating network, users understand that the map is not static and

they can move around and play with it, feature that they seem to mostly enjoy. Moving arranging, zooming, pinning etc are the network features which most augment users confidence and satisfaction with the tool.

On the contrary emotional reactions to the linear interface for arguemnts visualization regards general skepticism and annoyment which can also decay into confusion and feeling lost ("*I think I am lost*"-P1; *I am now buringn.. because I haven't work out how to do it*" "*that is the all page I am looking at...ohhh ok I give up!*"-P5). The main reasons of users frustration are related to the need of "*too many clicks*" to seek information and frequent "*change of context*" which often provokes disorientation.

CONTRIBUTIONS AND FUTURE RESEARCH

The grounded theory analysis of the experimentation's video and post experimentation interviews transcripts showed that different graphical interfaces for the representation of arguments strongly affect the way in which users read and understand the online discourse. Linear interfaces perform worse especially when when the information is nested into more articulated argumentation chains (Task 2). In such cases the network-like representation and the visual hints such as network structure, iconography and links' labels and colors seem to facilitate the identification of argumentation chains, thus supporting indirect connection and higher-level inferences of how the content connects. The results show that data model interpretation is also improved by argument visualization. Notably, exposing the data model in form of argument maps appears to enable a *learning by example* mechanism, whereby users reinforce their understanding of the data as they navigate through the user interface. Moreover the effectiveness of network visualization of arguments and the positive impact it has on arguments reading and comprehension increases as information complexity increases: the bigger the discussion, the better network visualizations performs compared to linear-threaded visualizations. Finally there is an element of fun and excitement associated to dynamic network visualization of arguments. This may suggest that this type or arguments visualization should be favored in fields such as gaming or education in which the element of "play" and positive emotional reaction are key factors to success.

Finally from a theoretical perspective we have presented a literature review of CSAV in the light of the late development of social media and Collective Intelligence research. This paper shows promising results on the capability of network visualization of arguments to support reading and sensemaking of online discourse activity, especially when information quantity and complexity increases, as for instance in the case of complex debates and massive online discussions. Network visualization of arguments also plays an important role in augmenting online discussion by providing a layer of structure that helps to improve human

comprehension of the argumentation structure behind the online discourse (by improving the understanding of the argumentation data model). These findings open new avenues for combining social media discourse with advanced network visualizations of discourse elements (issues, solutions and arguments) to both improve users' engagement and sensemaking of large-scale online deliberation processes.

REFERENCES

1. Brown, E. and Cairns, P., A grounded investigation of game immersion. In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, 2004). ACM. pp.1297-1300

2. Buckingham Shum, S., The Roots of Computer-Supported Argument Visualization. In: P. A. Kirschner, S. Buckingham Shum and C. Carr (Eds.), *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making.* Springer-Verlag, London, 2003.

3. Buckingham Shum, S., Cohere: Towards Web 2.0 Argumentation. In: *Proceedings of the 2nd International Conference on Computational Models of Argument* (28-30 May 2008, Toulouse, France. Preprint: http://oro.open.ac.uk/10421, 2008)

4. Buckingham Shum, S., MacLean, A., Bellotti, V. M. E. and Hammond, N. V., Graphical argumentation and design cognition. *Hum.-Comput. Interact.*, 12, 3, (1997), 267-300

5. De Liddo, A., Sándor, Á. and Buckingham Shum, S., Contested Collective Intelligence: Rationale, Technologies, and a Human-Machine Annotation Study. *Computer Supported Cooperative Work*, 21, 4-5, (2012), 417-448. http://oro.open.ac.uk/31052

6. DebateGraph: Web visualization of idea and debates. http://debategraph.org

8. Glaser, B. and Strauss, A., *Discovery of Grounded Theory Strategies for Qualitative Research.* Sociology Press, Mill Valley CA, 1967

9. Grasso, A. and Convertino, G., Collective Intelligence in Organizations: Tools and Studies (Eds. Special Issue). *Computer Supported Cooperative Work*, 21, 4-5, (2012), 357-473. http://rd.springer.com/journal/10606/21/4/page/1

10. Hill, W. C., Hollan, J. D., Wroblewski, D. and McCandless, T., Edit wear and read wear. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Monterey, California, USA, 1992). ACM. pp.3-9

11. Iandoli, L., Klein, M. and Zollo, G., Enabling on-line deliberation and collective decision-making through large-scale argumentation: a new approach to the design of an Internet-based mass collaboration platform. *International Journal of Decision Support System Technology*, 1, 1, (2009), 69-91

13. McCall, R. J., PHI: A Conceptual Foundation for Design Hypermedia. *Design Studies*, 12, 1, (1991), 30-41

14. Novak, J. D., *Learning, creating, and using knowledge: concept maps as facilitative tools in schools and corporations*. Lawrence Erlbaum Associates, Mahwah, NJ, 1998

15. Rahwan, I., Zablith, F. and Reed, C., Laying the Foundations for a World Wide Argument Web. *Artificial Intelligence*, 171, 10-15, (2007), 897-921

16. Selvin, A., Buckingham Shum, S. and Aakhus, M., The Practice Level in Participatory Design Rationale: Studying Practitioner Moves and Choices. *Human Technology: An Interdisciplinary Journal of Humans in ICT Environments*, 6, 1, (2010), 71-105. Eprint: http://www.humantechnology.jyu.fi

17. Sereno, B., Buckingham Shum, S. and Motta, E., ClaimSpotter: an environment to support sensemaking with knowledge triples. In: *Proceedings of the 10th international conference on Intelligent user interfaces* (San Diego, California, USA, 2005). ACM. pp.199-206

18. Shirky, C., *Here Comes Everybody: How Change Happens When People Come Together* Allen Lane, 2008