Making Online Discussions More Sense-able by Adding Richer Structure

Amy X. Zhang
MIT CSAIL
Cambridge, MA
axz@mit.edu

David Karger
MIT CSAIL
Cambridge, MA
karger@mit.edu

Abstract

As more discussions take place online, within applications such as forums, social media, chat, and email, many discussions have grown to be too large for an individual to reasonably handle. While tools for filtering and sorting may help users find and consume a particular subset of comments, there exists few tools for individual and collaborative sensemaking of large-scale discussion data, towards the eventual goal of gaining an overview of the discussion as a whole. In this piece, we argue for systems that allow users to mark up online discussions to create richer structures, using techniques such as annotation, tagging, and summarization. The addition of structure assists readers by providing signposts and places to dive in during the exploration of many threads. Such techniques can also scaffold higher level synthesis of the major points of the discussion, through the characterization of small amounts of discussion that then get recursively combined until the entire discussion is summarized.

Author Keywords

online discussion, forums, chat, social media, email, summarization, annotation, tagging, sensemaking

ACM Classification Keywords

H.5.3. [Group and Organization Interfaces]: Asynchronous interaction; Web-based interaction

Introduction

Visit almost any space online where people are conversing today, and it is easy to find examples of discussions that are overwhelmingly large. From pages of comments in Wikipedia Talk pages, to long streams of back-and-forth chat logs in Slack, to deeply threaded forums on news articles, massive online courses, message boards, and social media, it can sometimes feel like we are inundated with discussion while online. Encountering this much information, especially as a newcomer to the discussion, may lead to feelings of disorientation, with little indication of where to begin, how to navigate through the comment space, or keep track of deeply threaded replies, as well as feelings of overload, due to the unending steam of comments with no narrative or topical cohesion to make sense of their placement or to paint an overall picture. In many cases, the discussion grows so large that it is impossible for an individual to read the entirety of it—why then do interfaces choose to show all of it and with little guidance for sensemaking? These problems become exacerbated when we move from a single overwhelming discussion to many discussions of this nature, with oftentimes overlapping topics of discourse.

The abundance of massive online discussions has not always been the status quo. Online discussion has exploded in the past several decades in terms of size of conversation and number and diversity of participants. Large discussions online can now be found in many different domains as well, from health to education to civic discourse. Conversations can be anything from throwaway chitchat with few readers to deliberations with consequential ramifications and millions of views. So why do most discussion interfaces still treat discussions as dispensable data, with little affordance for organizing or synthesizing them? Indeed, online discussion systems and interfaces have changed little structurally since their inception, with almost all large discussions forming an unwieldy tree shape or never-ending linear chain of comments.

Given the challenges that people face with large-scale discussion today, we need new systems to provide users with the capability to better navigate through, organize, contextualize, and synthesize discussions. In this work, we present three categories of techniques for making sense of discussions by adding richer structure: tools to situate conversation in a particular context through the use of annotation, tools to support the characterization of threads and comments through the use of tagging, and tools to support synthesis of discussion into progressively higher level summaries using recursive summarization. The first two techniques assist with *foraging loops* during the sensemaking process, as introduced by Pirolli and Card [9], by providing points for readers to dive in to particular threads in the case of annotation, as well as providing more contextual cues to readers while navigating through comments and threads in the case of tagging. The third technique of recursive summarization is an instantiation of multiple *sensemaking loops*, where each iteration of summarization produces a higher level synthesis of the discussion, towards the creation of a hierarchical structure for exploration called a summary tree.

While the tools we envision could be used by an individual for their personal benefit, the collective use of such tools would allow sensemaking to scale with the size of discussion. Instead of requiring every person who arrives at a discussion thread to individually make sense of the discussion on their own as they must do now, these systems could harness greater numbers of readers and participants to build on each other's work.

Techniques for Making Sense of Online Discussions

We describe a number of techniques for adding richer structure to online discussion. We give examples of systems that use these techniques, oftentimes combining more than one, to support sensemaking.

Annotation

The first set of systems anchor conversations to parts of a separate primary piece of content through the use of annotations “in the margins". This technique creates a structure where a piece of text contains separate, smaller discussion threads dangling off of it in various locations. The benefit that this provides is the ability to separate out disparate discussions and place them in context of what they discuss, so that discussions near each other in space are about the same thing. Another benefit this provides is the enrichment of a primary piece of content, whether that be a webpage, a news article, or a textbook, that can serve as a launchpad for diving into discussions. While scanning a document, readers can be clued in on specific passages or pages that generate a great deal of discussion, through a visual representation of the density of annotations, and from there, explore the comments. Systems of this variety include Nota Bene (NB) [13], for discussion anchored to passages within educational materials. Other systems that support this feature include Hypothes.is, Fermat’s Library, and blogs such as Medium.

Tagging

The second set of systems allow for tagging of individual comments or subthreads with relevant contextual information. Depending on the nature of the tags, this information could help separate out comments within discussions into topical or other categories for the purpose of organization or as a filtering mechanism, or provide greater contextual understanding of the discussion, such as by annotating the veracity of claims or the underlying framing of arguments.

Much like in other information sensemaking tasks, the process of tagging can assist in emergent taxonomy creation towards higher level sensemaking. The Wikum system [12] for summarizing large discussions incorporates open-ended tagging and grouping by tag to make it easier to summarize similar comments. Certain types of tags can also serve as signposts for readers while navigating through many threads of replies or long chains. An example of this is also in the Wikum system, which allows users to attach a summary to a subthread so that future readers can get a quick sense of the thread before deciding whether to dive in. Another example in the realm of linear chat streams is in a prototype tool called Tilda, which allows the annotation of major discourse acts, such as “action item” or “announcement” within a chat log using a taxonomy of emoji reactions to provide more signals while scrolling through a backlog of chat messages. A place where tagging is used in the wild is Reddit’s ChangeMyView forum, where debaters mark when someone’s argument has succeeded in changing their mind.

While tags of comments and threads on their own do not require alternative structures of discussion, their combination with other techniques can result in richer structures. One example is a feature in NB that allows students to indicate their emotion expressed in a comment, whether it be confusion or curiosity [11]. The combination of annotation and tagging allows for the primary text to now visualize the emotional reaction of students to different parts of the material. This is helpful for students to determine what to read as well as teachers to determine what broader concepts needs to be addressed when faced with a sea of comments. Another example of where this happens is in

law schools, where students often highlight their case studies in different color markers to differentiate different parts of the case.

Summarization
The final technique helps make sense of large discussions by providing higher level summarization capabilities within discussion systems. These systems provide a way for people to reduce scale and redundancy as well as a way for people to reflect on a holistic understanding of what was said. Examples of places where summarization is used to synthesize discussion include Quora Answer Wikis, a collaborative summary of all answers to a question, and Wikipedia Talk Pages, where deliberations are often officially "closed" and summarized by a neutral party. But writing a summary of a large discussion is a massive task, unlikely to appeal to the many readers who do not even bother to read the entire discussion. Indeed, as we've found in interviews with frequent Wikipedia closers, it may take hours to read a single discussion and write a closing statement. To address this problem, we consider how summarization of a large discussion can be decomposed into small amounts of work that can be done while exploring a discussion.

The Wikum system facilitates this process by scaffolding a recursive summarization workflow, where users build summaries of small subthreads of discussion, small sets of those summaries are then aggregated and summarized, and so on, until the entire discussion is summarized. The resulting structure is that of a dynamic textual summary, called a summary tree, that can be explored at varying levels of detail. By building higher-level summarization on top of lower-level tagging and signposting of subthreads, Wikum guides an individual or group through the sense-making process. From studies we have conducted, we have found that newcomers retain more information from exploring a summary tree rather than a linear document of the discussion with inline summaries. The Tilda prototype provides a similar functionality but in chat conversations. Tilda supports the grouping of chat messages into higher level conversation blocks, where the aforementioned tags become "meeting notes" or main takeaways for that conversation. The conversation notes can then be referenced in different places in chat, grouped into broader themes, or summarized further in other knowledge management applications.

Comparison to Complementary Methods

There are a number of ways to manage unwieldy online discussions, some of which could be used in tandem with the methods we present, while others have downsides to consider.

Social Moderation and Collaborative Filtering
Most existing systems now incorporate some form of collaborative voting process to sort and filter comments. However, there are documented problems, including underprovision [4] and negative feedback loops [2]. Social moderation may surface only "popular" points and push down minority opinions. Additionally, filtering techniques can only go so far. Comments may still be too numerous, have many tangents, and be redundant. Better mechanisms for personalization or recommendation can filter down some of the noise but also may lead to "filter bubbles" when only one point-of-view is represented.

Imposing Structure at the Outset of Discussion
Some specialized online discussion systems impose a particular structure of discussion at the outset. One case of this is community question-answering (CQA) forums such as Quora or its predecessor Answer Garden [1], which require all discussion to be in a question-and-answer format.

Other systems restrict users to creating pro and con lists, such as ConsiderIt [7], or only allow users to respond to points with support or opposition, such as Kialo. Finally, other systems such as Arguman and the Deliberatorium [6] expand the range of discourse acts to turn conversations into argumentation maps, requiring respondents to conform to one of several possible types of arguments. NB and some other annotation systems also force participants to anchor their comments.

The benefit of this approach is that the enforcement of structure provides a shortcut towards making sense of the discussion. Instead of creating structure through the use of techniques like tagging, discussions are already structured to a degree, making it easier to perform higher level synthesis. There are several downsides to this approach though. One is that the enforcement of structure at the outset can lead to changes in the type of discussion that emerges or outright restrictions in the type of discussion possible. For instance, CQA sites work well for questions that have a clear "best answer" but not as much for questions where more answers provide more information, such as in the case of opinion-seeking questions or requests for anecdotes [8]. By enforcing certain structures, some systems may also remove the "fun" out of participating in and reading discussions. For instance, the debate-like style of some structured discussion systems may be forcing out witty banter and personal anecdotes and replacing them with a drier set of claims and counterclaims. In contrast, systems like Wikum allow the discussion to happen naturally and then builds structure in after the fact.

Implicit Signals and Modeling

Some systems seek to mine discussion data or analyze implicit signals in user behavior to extract information to help in sensemaking. Examples of this include discourse act modeling to predict the discourse structure of discussions or stance classification of comments [10]. These kinds of models could be helpful for more easily constructing argumentation structures without forcing them upon comment writers at the outset. However, while models trained on implicit signals reduce the need for users to contribute work, they oftentimes are not sophisticated enough to understand the language in conversations. There may be important context that is not explicitly expressed in the text of the comment as well. Still, there are cases where such models could assist users in their sensemaking, such as by suggesting possible tags for a comment or highlighting key sentences that might form a summary, which are both features of Wikum. Another example is ConVisit [5], which uses a combination of user inputs via tags and topic modeling to generate a visualization of the major topics within a discussion.

Visualizations

Finally, many systems convert explicit or implicit signals into a visual representation of the discussion. One example of this is OpinionSpace [3], which maps comments onto a 2-D display so that readers can explore with an eye towards finding similar or diverse opinions. Visual representations of discussions can often be helpful for sensemaking as they provide an overall picture of the discussion and places to dive in. Indeed, both NB and Wikum provide visual cues. One drawback is when visualizations are too abstract, such as graph-like diagrams with nodes for comments, they can feel too foreign to a certain subset of readers or too complicated for casual readers. We noticed this with some users who explored a visual representation of Wikum summary trees. Another drawback is that large visualizations are difficult to represent in smaller devices such as in mobile. These reasons may be why interactive visualizations of comments and threads are not often present in the wild.

Discussion

Broadening out, both the structure provided by annotation and the structure of a summary tree could be seen as forms of restricted hypertext, where instead of being able to navigate through a graph of arbitrary documents, the structure is instead hierarchical. That is, it is only possible to move up and down, or zoom out and zoom in to discussion at varying levels while navigating systems like NB or Wikum. But by being slightly more restrictive than hypertext, perhaps such a structure reduces the feeling of being "lost in hyperspace" that can be accompanied with non-linear reading on the web.

Broadening beyond discussions, the techniques we describe can also be applied to arbitrary text, and the structures can lend themselves to authoring tools. For instance, one could imagine a recursive summarization approach to writing a paper, where different levels of the hierarchy present different levels of detail, much like how a title summarizes an abstract which summarizes an introduction, and so on. Alternatively, a textbook could be written with pre-authored annotations in the margins or "sidenotes", much like how footnotes operate today.

Conclusion

As discussions online continue to grow, we need systems to help users make sense of them. We present three techniques to individually or collectively build richer structures of discussion towards the goal of better exploration and gaining an overview. We compare our approach of enabling user actions to anchor, mark up, or summarize discussion with common approaches such as voting and structuring discussions at the outset. We also discuss how the collection of implicit signals and learned models as well as the use of visualizations could be used in tandem with such techniques.

REFERENCES

1. Mark S Ackerman and Thomas W Malone. 1990. *Answer Garden: A tool for growing organizational memory*. Vol. 11. ACM.
2. Justin Cheng, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2014. How community feedback shapes user behavior. *arXiv preprint arXiv:1405.1429* (2014).
3. Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1175–1184.
4. Eric Gilbert. 2013. Widespread underprovision on Reddit. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 803–808.
5. Enamul Hoque and Giuseppe Carenini. 2015. Convisit: Interactive topic modeling for exploring asynchronous online conversations. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, 169–180.
6. Mark Klein. 2011. How to harvest collective wisdom on complex problems: An introduction to the mit deliberatorium. *Center for Collective Intelligence working paper* (2011).
7. Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 265–274.
8. Lena Mamykina, Drashko Nakikj, and Noemie Elhadad. 2015. Collective Sensemaking in Online Health Forums. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3217–3226. DOI: http://dx.doi.org/10.1145/2702123.2702566
9. Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. 2–4.
10. Amy X. Zhang, Bryan Culbertson, and Praveen Paritosh. 2017a. Characterizing Online Discussion Using Coarse Discourse Sequences. In *Proceedings of the 11th International AAAI Conference on Weblogs and Social Media (ICWSM '17)*.
11. Amy X. Zhang, Michele Igo, Marc Facciotti, and David Karger. 2017b. Using Student Annotated Hashtags and Emojis to Collect Nuanced Affective States. In *Proceedings of the Fourth Annual ACM Conference on Learning at Scale: Poster Publication (L@S '17)*. ACM.
12. Amy X. Zhang, Lea Verou, and David Karger. 2017c. Wikum: Bridging Discussion Forums and Wikis Using Recursive Summarization. In *Proceedings of the 20th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '17)*. ACM.
13. Sacha Zyto, David Karger, Mark Ackerman, and Sanjoy Mahajan. 2012. Successful classroom deployment of a social document annotation system. In *Proceedings of the sigchi conference on human factors in computing systems*. ACM, 1883–1892.