

A Tagging Approach for Bundling Annotations

Yamin Htun, Joanna McGrenere, Kellogg S. Booth
Department of Computer Science, University of British Columbia

{yhtun, joanna, ksbooth}@cs.ubc.ca

ABSTRACT

In a paper presented at CHI 2006 we introduced structured annotations, called bundles, to support co-authors in the edit-review-comment document lifecycle, and we reported a study showing that bundles facilitate workflow by improving reviewing accuracy and efficiency. Bundles are a “top down” way to organize annotations. We demonstrate an enhanced prototype that also supports “bottom up” organization using tagging techniques, new automated bundle creation options, and the reviewing features and manual bundle creation present in the first prototype.

Categories and Subject Descriptors

H.5.3. [Information interfaces and Presentation, HCI] Group and Organizational Interfaces – *Asynchronous interaction, Computer-supported collaborative work*

General Terms

Design, human factors

Keywords

Asynchronous collaboration, collaborative writing, tagging, structured annotation

EXTENDED ABSTRACT

Asynchronous collaborative writing is common, and annotations play an important role as a central communication medium connecting co-authors with evolving artifacts in the process [7]. However, the lack of support for rich annotations in most word processing systems often forces valuable communication to happen outside the shared document in the bodies of emails, to which the document is an attachment. These messages are separate from the document, making the establishment of a shared reference for discussion difficult [2].

Co-authors often copy and paste referenced content of the document into email or type explicit navigation statements such as “Clarify my questions on the third and last paragraphs,” which can be time consuming and error-prone. Significant overhead is required to reconstruct the context of the communication [4]: workflow requires navigating between email messages and the document itself [4] and information is likely to be lost or ignored [1]. At best, in order to keep track of the workflow and progress in the task, collaborators need to maintain not only document files but also the email messages [8]. Information overload and

workflow inefficiencies can result with increasing numbers of annotations after only a few reviewing cycles.

To facilitate the workflow management involved in collaborative writing, we previously identified user-centered requirements for annotation support and developed a comprehensive model of annotations [8] in which each annotation has a set of attributes such as the *creator* of the annotation, a *timestamp*, reviewing *status* (read/unread and accepted/rejected), and one or more *anchors* to material in the document. Annotations can have optional attributes such as a list of *recipients*, a *comment*, *replacements* for the anchored material, a *name*, and *substructure*. A *bundled annotation* (or *bundle*) represents a structured group of annotations with various anchors into the document. There are no restrictions on structuring annotations other than that they be acyclic; an annotation can be associated with more than one bundle. Changes in an annotation’s status will be automatically synchronized across different bundles to which it belongs.

We previously described a user study that investigated the effect of structured annotations on reviewing workload and quality [8]. Participants were asked to review a set of annotations with a Simple Editor containing only basic annotations (edits and comments) with high-level communication taking place in a separate email message window, and with a Bundle Editor in which annotations are structured into bundles with high-level communication integrated as generalized annotations. Participants performed faster and more accurately with the Bundle Editor and they found bundles innovative and intuitive. We did not investigate the usability and consequences of bundles in the annotation-creation stage. We are now examining this.

In our model, bundles can be created in four ways: (1) *manually*, (2) *automatically*, (3) as a result of *filtering operations* and *queries*, and (4) as a result of *editing commands*. While annotating the document, co-authors *manually* create bundles by explicitly selecting and grouping annotations into bundles. At the end of each reviewing session, a bundle is created *automatically* with all the new annotations made during the session. Every time a user *filters* the annotations based on specified attributes, a temporary bundle is created, which can be saved as a permanent bundle with a single click. Moreover, when a user performs normal *editing commands* such as “Find/Replace” or “Spell Check”, a bundle will be created with all the edits from the command gathered into sub-bundles such as “replaced,” “skipped,” and “ignored”.

Although automatic bundle creation does not require extra effort from reviewers, we doubt that automation can fully capture the richness and complexity of the annotations used in discussions. Hence, our goal is to minimize the effort required by reviewers when manually creating bundles and managing annotations. While exploring different approaches we were inspired by recent successes with *tagging*, in which users assign meta-data or keywords to information resources. Traditionally meta-data is created by professionals (catalogers or authors) [5], but systems

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW 2006, November 4-8, 2006, Banff, Alberta, Canada.

Copyright 2006 ACM ...\$5.00.

like **flickr** and **delicious** allow ordinary users to describe and organize content with any vocabulary they choose. Tagging facilitates the organization of information within personal or shared information spaces. Browsing and searching tags attached to information resources by other users encourage collaboration. Compared to traditional folder-based hierarchical information management models, collaborative tagging is believed to reduce the cognitive workload experienced by users [6]. A major drawback for tagging is the ambiguity and imprecision of tags and the lack of control for synonyms and homonyms [3].

In our top-down approach, a user associates an annotation with a bundle by manually dragging the annotation into the bundle. When an annotation is in multiple bundles the work increases linearly with the number of associated bundles. Tagging is a bottom-up approach that reduces effort and achieves a more seamless workflow. An annotation can be easily associated with more than one bundle simply by tagging it with appropriate keywords; bundles are created through filtering that recognizes tags as filterable attributes. Because co-authors have their document as a shared context, we believe tags will be consistent and scalable across users, alleviating the ambiguity and imprecision seen in more general contexts while providing flexibility in classifying information into more than one category. Bottom-up tagging captures multiple semantic concepts that are inherent in most information resources through a light-weight and intuitive means of organizing and sharing information in a collaborative setting.

The core interface to the “Bundle Editor” prototype consists of a *document pane* and a *reviewing pane* (Figure 1). The main component of the document pane is the document editor, which has typical functionality (insert, delete, comment, etc.). The reviewing pane is a multi-tabbed pane with each tab displaying a specific group of annotations. The reviewing pane supports creating new bundles, adding and removing annotations from a specific bundle, and sorting and filtering annotations based on

particular attributes.

Tagging, which is bottom up, is appropriate for unknown workflows where structure emerges and serendipity needs to be supported. During more precise workflow, top-down structuring through manual or automated bundle creation is likely to be the preferred approach. We will demonstrate both top-down and bottom-up structuring in the Bundle Editor to illustrate the advantages of each. We expect to report results from preliminary studies of how co-authors use these two approaches. The studies will compare ease of use across the two approaches, examine the semantic categories within annotations for a shared document, and investigate the role of bundles in facilitating problem decomposition strategies involved in co-authoring workflow.

REFERENCES

- [1] Cadiz, J., Gupta, A., Grudin, J. (2000). Using web annotations for asynchronous collaboration around documents. *ACM CSCW '00*. pp 309-318.
- [2] Churchill, E., Trevor, J., Bly, S., Nelson, L., and Cubranic, D. (2000). Anchored conversations: chatting in the context of a document. *ACM CHI '00*. pp 454-461.
- [3] Guy, M., and Tonkin, E. (2006) Folksonomies: Tidying up Tags? *D-Lib Magazine*, 12, 1.
- [4] Hee-Cheol, K., and Eklundh K. (2001). Reviewing practices in collaborative writing. In *Computer Supported Cooperative Work*, 10, 2. pp 247-259.
- [5] Mathes, A. (2004) Folksonomies – cooperative classification and communication through shared metadata. <http://www.folksonomies.com/academic/computer-mediated-communication/folksonomies.html> (accessed 07/2006).
- [6] Sinha, R. (2005). A cognitive analysis of tagging. In *Rashmi Sinha's weblog* http://www.rashmisinha.com/archives/05_09/tagging-cognitive.html (accessed 07/2006).
- [7] Weng, C., & Gennari, J. (2004). Asynchronous collaborative writing through annotations. *ACM CSCW '04*. pp 578–581.
- [8] Zheng, Q., Booth, K.S., and McGrenere, J. (2006). Co-authoring with structured annotations. *ACM CHI '06*. pp 131-140.

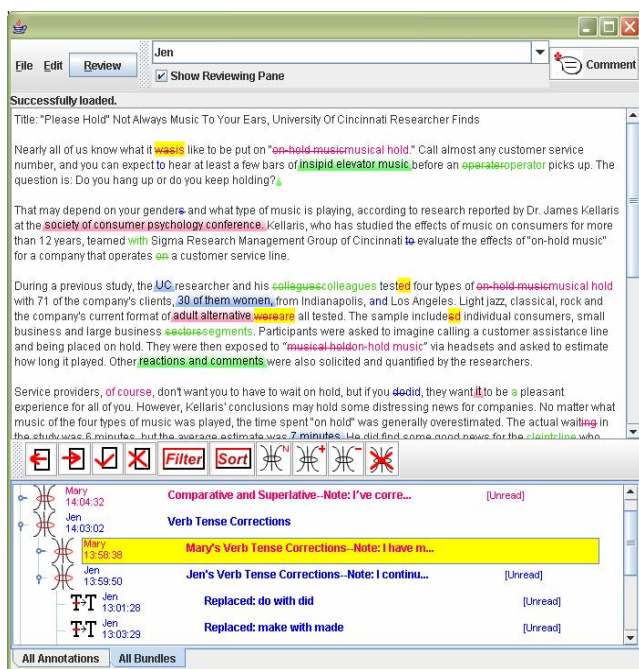


Figure 1. Bundle Editor with document and reviewing panes.