

Conversation around documents: more than threading

Henry Rodriguez

IPLab/NADA The Royal Institute of Technology
of Sweden
NADA KTH 10044 Stockholm
henry@nada.kth.se

Kerstin Severinson Eklundh

IPLab/NADA The Royal Institute of Technology
of Sweden
NADA KTH 10044 Stockholm
kse@nada.kth.se

INTRODUCTION

Besides being a great repository of data, the Web is a space for discussion. More and more systems that aim to support distributed group activities are being developed e.g., Yahoo groups and Google groups. Most of these systems offer an option that allows their members posting messages and therefore discussing different topics. These discussions appear mainly in the form of text, they are carried out asynchronously, and they are persistent in time. Furthermore, the topics of these discussions are very diverse and they need to be sorted out.

A crucial design issue is how these messages are presented to the users, helping them to understand the relationship among the messages. A very simple and widely mechanism that has been used to chunk these conversations is by emphasizing the reply-to relationship between messages by structural and visual means. This mechanism is known as threading. However, the reply-to relationship cannot thoroughly depict the complexity of a conversation. A message can be related to a previous message although it is not necessarily a reply to that message.

How users become aware of the current activity of an asynchronous, distributed, online, text-based conversation is another concern when designing such systems. A common mechanism used is to highlight the longevity of the messages in the conversation, e.g., *what is new*. Another mechanism that has been used is by sending an email notification to participants for announcing that new comments have arrived. The notification may be sent e.g. to one person e.g., the moderator, or to each member of the group. However, selecting to whom the notification will be sent is not in the hand of the sender and it is usually not taken into account how interested a person is in receiving these notifications.

Most of the information presented on the Web is in the form of documents. It is natural that users would be interested in discussing a set of documents. However, little support is given to users who are interested in getting involved in a conversation about a set of documents. To discuss around documents implies that the structure of the conversation can be more complex. The thread of the conversation involves not only the messages that have been posted but also the documents that are being discussed.

In the light of these issues, we present an alternative design for discussing a set of shared Web-documents, presenting the messages of these persistent conversations to users, and making people aware of the ongoing discussion. These design alternatives have been implemented in a Web-based system that we have developed. The system and the design alternatives will be presented in the next section.

COL-LECCIÓ AND THE NOTIFICATION OF NEW COMMENTS

Col-lecció is a system that supports asynchronous conversations around Web-documents in small or medium-sized groups. The collection of documents and their annotations belonging to a common task managed by Col-lecció is what we call a *domain*. To add a document to the system, users only need to give the document's URL. Col-lecció presents the documents using their titles forming an index of content. To select a document, users just click to the corresponding title. When a document is selected, Col-lecció shows its content, and the annotations made to it. The annotations are presented in a chronological order and no threading mechanisms are used.

To add a new comment, the system pops-up a new window, the Add Comment Window (ACW). This window presents two spaces. The left space is the context-reconstruction part. This space presents the document content and the comments that have been made so far in a single flow. The comment itself can be written in the right space or the input area of the ACW. When the comment is added, an email notification is sent automatically. In previous studies, [3], we have observed that the email notification was appreciated by users. Many of them indicated that the email notification was like a trigger or a reminder of the task. However, users also indicated that this could increase their email traffic (which was already considered as high by many of the users).

Email notifications can actually be a way to affect how conversations develop. To whom the email notification will be sent, is a concern that we have surveyed while developing Col-lecció. From our investigations and evaluations, we have introduced a novel concept: a Growing Flexible Mailing List (GFML). A GFML is a check list used to involve users in a conversation and to decide to whom the notification is sent.

The input area of the ACW presents a GFML. Each item of the check list is formed by names of participants. The initial

item in the GFML is the user who added the document to the system. When a comment is added, the system checks if the user who is adding the comment has participated in the conversation earlier. If not, the user is added to the check list. Thus, the check list grows. We assume that if a person has commented on a document, this person is interested in the document and might want to be coupled to that conversation. The email notification will be sent to those who were checked-on (selected) from the check list. By default, the system checks on all items in the list.

The process of selection is simple and flexible. For example, Louis, a participant who is commenting on a document that he has commented before, might uncheck his name from the check list to avoid sending the message to himself. Additionally, by looking at the check list, users can know how many and who have participated in the discussion of a particular document (see figure 1). We argue that a GFML can affect the way a conversation is carried out. For example, making comments on many documents means that the users might be involved in many conversations and therefore, his/her email traffic is increased. This might hinder users from making more comments on those documents that he/she has not commented. On the other hand, it could be known which users have made no comments in the domain and this might cause them to comment on a document. There are still many issues to investigate with regard to GFML. For example, how would participation change if all items are automatically checked-off by default?

THE REFERENTIAL SPACE IN A DOMAIN

Another interesting aspect in relation to conversations around documents is the referential space in which the comments are presented. Observe that in a domain there can be as many conversations as the number of documents in it. Comments in a domain can mainly refer to:

- In-situ text
 - What has been said in the conversation (a previous comment)
 - The content of the document in discussion.
- Off-situ text:
 - What has been said in another conversation in the domain
 - The content of another document in the domain.
- Text outside the domain, e.g. other Web documents.

Discussions may also focus on the system itself, which is given a separate space in the system. Finally, our studies have shown that comments are often made about the ongoing discussion on a meta-level. The extent to which the above kinds of references actually occur depends on the particular context of the dialogues: the purpose and nature of the collaboration; what aspects of the document are being negotiated, etc. (see [4] for a discussion).

VISUALIZATION OF THE CONVERSATION IN A DOMAIN

We have created a visualization tool, DID (Domain Interactivity Diagram) [4] to help clarify the conversational activity exhibited in Col-lecció. The purposes have been partially explorative: to assess the interactivity and referential relationships within a particular domain and as a tool for comparison between different discussions in this regard. However, we believe that visualizations like the ones we are using might also be helpful to participants themselves as optional coherence tools. This would increase the accessibility of the discussion space, and thereby its “social translucence” [1], i.e. the degree to which other people's activities are immediately available and visible to a user.

The aim of the visualization of a domain is to help participants to get an overview of the general character of a discussion at a glance, before reading. For example: Which was the document that was most commented on in the domain? Was the document discussed by many of the participants? Did the author of the document participate in the discussion? The answer to these questions might help the user to choose her reading strategy. Moreover, a participant could assess a comment according to what is known about the comment's author in relation to the task. For example, knowing whether a participant has read the other documents in the domain or not, would help as to value his/her point of view about the document.

The DID presents a domain at two levels. The first level is a general overview of the domain; the *panorama view* (see figure 2), and the second level focuses on the discussion of a single document; the *document interactivity view* (see figure 3). Common to both views is that each participant in a domain is assigned a color. Objects created by him or her, that is comments or documents, are presented in the same color. With this information it is easy to see how each person has spread out his/her activities among the different discussions. As a social navigation feature this may both support coherence and promote the group's discussion when the number of participants increases.

NO THREADING

An important issue is how traditional threading would affect communication in Col-lecció. The questionnaires and interviews with users showed that threading is often a desired feature. We avoided threading in the design of Col-lecció, mainly because the discussions as a whole were expected to focus on the shared document, and we expected that threading might contribute to the topic drifting away from the shared document. Separating threads can also impose an additional burden or cost on the user, who has to indicate to which thread a contribution belongs [2]. In summary, we believe that the nature of a text-based conversation around a set of documents differs from other form of CMC. Col-lecció, DID and the GMFL are interesting design alternatives that deserve further investigation.

