WikiPlayer: Press play to start the wiki evolution

Johann Larusson
Department of Computer Science
Brandeis University
415 South Street, Waltham, MA 02454
johann@cs.brandeis.edu
+1 781-736-2712

Richard Alterman
Department of Computer Science
Brandeis University
415 South Street, Waltham MA 02454
alterman@cs.brandeis.edu
+1 781-736-2703

ABSTRACT
In recent years "wiki" web sites have enabled people to collaborate online by allowing content to be modified freely by any user. This platform has proven to be very successful despite its unconventional "anarchistic" and chaotic organization – Wikipedia being the most famous example.

When people contribute to a wiki web site their work takes place within a representational system comprised of multiple distributed representations such as the wiki pages and external web sites related to the wiki topic. How much of the collaborative work involves transferring information between the different representations? How does the collaborative work emerge within the system and how is it used by the participants to accomplish their goal?

We present the WikiPlayer, a tool to visualize and replay the entire revision history of related wiki pages as they collectively evolve over time. The player allows us to track each user's contribution to a set of wiki pages, review the state of each page at any given moment in the history, and easily generate statistics helpful in analyzing the collaborative community of practice. The tool can be used to identify the collaborative work patterns that develop from the emergent interaction between the structure of wiki pages and the organization of the participants' representational work.

Author Keywords
Wiki, revision history, replay, collaboration, representational system.

ACM Classification Keywords
H.5.2: GUI; H.5.3: Collaborative computing, Web-based interaction.

INTRODUCTION
Collaboration is the process of interaction amongst people who share the same goal [2]. Examples of collaborations include non co-located employees of the same company who need to work together on a project or college students that are required to collaborate on class assignments. Often collaborators are faced with daunting tasks such as agreeing on a specific date and time where everyone can meet, sharing information between participants, and viewing others' contribution which can be especially difficult within a large group using many sources of information. Development of online collaborative groupware aims at solving or simplifying these issues by supporting *different place* collaboration. One example of such a technology is a wiki invented by Ward Cunningham [13, 8] in 1995. A wiki is a special kind of web site that supports asynchronous editing and has two distinguishing main features. First, every visitor to the site has the rights to edit everything, and editing is not discouraged but encouraged. Wikis enable collaborative building of web pages by allowing people to freely edit both the content and the structure. Second, most wikis make public all past revisions of the articles that the users create, providing them as well as researchers with a rich record of the collaborative activity.

Traditionally the unit of analysis for cognitive science has been the individual. The *distributed cognition* (DCOG) view is that cognition is embedded into a large system of internal and external distributed representations. Thus, when people collaborate online using a wiki their collaboration is mediated by the wiki in terms of the representations available to the collaborators. Each individual has their own interpretation of the activity and of the content stored on external representations, such as books or remote websites, which contain information relevant to their collaborative project. The wiki pages that the collaborators create and maintain are a significant component of the representational system in which the collaborators work. It is for these reasons, that an analysis of cognitive performance (the collaboration on the wiki) can be framed in terms of the entire representational system in which the behavior is embedded [3, 4].

In this paper we describe our investigation into how collaboration on wiki sites is carried out with respect to the representational system. The WikiPlayer tool presented in this paper enables visualization and replay of a revision history of wiki pages. More than one wiki page can be replayed at the same time simplifying the process of keeping track of user contributions between the different representations (wiki pages). Since wikis make available the entire revision history of a collaborative wiki project the data set can be enormous. Data mining on such a vast set of information is not a straightforward task, let alone trying to analyze the data in terms of the entire representational system – the WikiPlayer addresses these issues.

MOTIVATION

At the Department of Computer Science we have been developing groupware applications that facilitate online collaboration and analysis tools to examine the collaborative activities by replaying complete transcripts of online practice [5]. The revision history provided by wikis is a partial transcript of the online collaboration mediated by the wiki system; a complete transcript, of the sort produced by our homegrown applications, enables the analyst to replay all online activity as if he or she were viewing a VCR tape. Building analysis tools, such as the WikiPlayer, that are capable of replaying transcripts of online collaboration is beneficial to both our research and teaching agenda.

We have introduced wikis for several educational and research related tasks. We have used wikis as a research platform to collect data of online collaborative activities to support our studies of synchronous and asynchronous collaboration. In the spring of 2006, one of the assignments in Internet & Society, a course taught at Brandeis University, required the class to collaboratively write a paper on one of the books from the reading list using a wiki as a collaborative platform. In another class, Computational Cognitive Science, wikis were used to collect replayable data of online collaboration both to support lectures and to use as a source of data for student term projects. For each of these wiki-based tasks we have developed several techniques for reviewing transcripts of the online community of practice.

One of our analysis tools replays transcripts produced by our groupware in such a way that it provides the analysts with an "over-the-shoulder" view into the online collaboration just as if they were replaying videotapes of the activity in a VCR [5]. Another vehicle of analysis we have developed enables analysts to extract discourse from the transcripts and view the conversational part of the collaboration as a hierarchically organized text providing also detailed metadata if desired. In this paper we discuss the WikiPlayer, a tool for visualizing collaborative work patterns that emerge as a product of representational work within a representational system where the collaboration takes place. The WikiPlayer enables the analyst to replay the flow of information as it is added and modified on the wiki.

In an educational context the WikiPlayer can make several valuable contributions to both students and teachers. Teachers can easily evaluate individual student contributions in a collaborative wiki assignment such as the one described above. The WikiPlayer can be used by students to better understand and keep track of each other's work during class projects as well as providing them with a in-depth view into the activities when studying analysis, design, and cognitive engineering methods that apply to online collaboration.

As a research tool, the WikiPlayer facilitates researchers in identifying collaborative work patterns and track the online collaborative activity and how it is affected by the representational system. The WikiPlayer enables close examination into how participants organize their representations, the propagation of representations between alternate forms, and the division of representational work. Some of the features of the WikiPlayer are fully automatic while others currently require brief additional "by-hand" analysis. However, as future work we intend to automate those features as well as add new powerful analysis functionality.

We have already used the WikiPlayer for several different tasks. In this paper, we will look at the latter function - a tool for conducting research on online collaborative activities. In this capacity we can demonstrate some of the more interesting functionality, and potential new features, of the tool.

We conducted an experiment with undergraduate and graduate students where they collaboratively built wiki pages describing a class schedule for an incoming freshman. The wiki used for the experiment was MoinMoin [12] one of many clones of the original wiki developed by Ward Cunningham. MoinMoin shares some of the main features with other wiki clones such as allowing all users to edit the content anonymously, and providing the revision history of the wiki pages.

Our main results show that there is a close relationship between the success of the collaborative project and the amount and kind of representational work carried out by the participants. The success of the activity is dependent upon how the participants recognize, share, and manage representations vis-à-vis the emergent representational system. We use the WikiPlayer to identify the kinds of representations and work that each team of users perform. Our results show that how a group organizes their representational work is predictive of how well they achieve their cooperative aim. In itself this result is not surprising. What is of value is that we can use the WikiPlayer to construct detailed models of how each group works together on their collaborative representational task.

WIKI TECHNOLOGY

A wiki is a freely expandable collection of interlinked web pages. In essence it is a database of documents that can be easily edited by anyone using an editor accessible through most web browsers [8]. People collaboratively write documents in a very simple markup language where they can either edit an already existing page or create new pages. Wiki pages, as presented to each visitor upon his or her first visit to the wiki site, are more or less just plain HyperText Markup Language (HTML) pages. The contributing user never edits the HTML code. Instead, using a simple editor the user edits the source of the wiki page, often referred to as the wikitext, which can also be augmented with a

Figure 1. Recent changes on MoinMoin wiki.

simplified markup language to indicate various structural and visual conventions [11]: wikis provide a simplified process of creating HTML pages. To contribute to a wiki topic users simply click on the "edit" button of the wiki page they wish to edit which will display the editable wikitext inside an editor.

Revision history: Tracking page edits

Most wikis, such as MoinMoin, include various tools that focus on providing users with a simple way of monitoring the state of the wiki as it is constantly changing. One such feature is the archival system. Wikis record and log every change made to a page no matter how insignificant the change might be. An important part of online collaboration on a wiki is keeping track of edits made to a page. The most recent changes to the wiki are accessible by clicking on a "Recent Changes" link that will open up a page that lists all the recent changes in a simple manner [Figure 1]. This functionality helps users keep track of the evolution of the topic they are interested in and monitor the contributions of their fellow collaborators.

One of the features provided by MoinMoin and most other wiki clones is the availability of raw data. Wikis make accessible the entire revision history of all documents created on its servers, a valuable source of information for researchers analyzing online collaborative activities. The archive can encompass tens, hundreds, or even thousands of revision entries for a single page. The log that the MoinMoin wiki provides contains information such as the name and/or internet address of the contributing user, the time and date when the revision was made and a reference to the actual text file containing the revision.

ANALYZING THE REVISION HISTORY

In the previous literature researchers have studied collaboration via wiki by examining the page revision history. Viegas et al. focused on visualizing the revision history of a single page to identify relationships between multiple document versions in order to reveal patterns within the wiki context that could be useful in other collaborative situations as well. While their analysis is interesting, and their results valuable for wikis and collaborative software in general, their methods do not take into consideration the evolution of the collaboration in terms of the representational system. Hutchins has produced detailed models of the representational systems for both the airplane cockpit and the bridge of a navy vessel [3,4]. When collaboration occurs within a representational system, work patterns emerge as a result of representational work required on behalf of the participants to carry out and complete their task [6].

A wiki topic can consist of multiple wiki pages where each page spans a specific subset of the topic. We refer to this as a *wiki project*. This organization is frequently observed on sites such as Wikipedia. The wiki project in addition to other external representations such as remote web pages used as information sources create the representational system. The collaboration occurs within this system and the wiki revision history describes the process of the collaborative activity. Therefore, when analyzing online collaboration it is important to look at the emergence of the collaborative wiki project as a whole and not just the evolution of a single page. The revision history archive can be quite large, even for only few pages especially when including all revisions. As an example the data we collected included roughly 2450 files, describing about 500 unique revisions to the wiki pages and totaling 2.4MB of raw text. Making sense of the history of the pages that the data set describes is challenging enough, analyzing the relationship between them at the same time is not straightforward.

WIKIPLAYER

The MoinMoin wiki provides an option of viewing the changes that have been made recently to pages on the server. This functionality is similar to what other wikis provide, in that it enables users to easily keep track of the constantly updating wiki, and see what documents were recently changed, who changed them, and when. However, there are two problems with the design of this interface. First, it is not possible to easily observe the evolution of related pages at the same time. Second, this only allows us to see recent changes and not the entire history. The *history flow* visualization technique [10] addresses one of these issues but has one major drawback, it only enables researchers to easily track trends in the collaboration by viewing the revision history a single page at a time. This does not take into consideration the importance of viewing the collaborative work effort by the community in terms of the representational system and how it affects the activity.

As discussed above, the WikiPlayer has several significant functionalities. The focus of the paper will be on how the WikiPlayer can be used to simultaneously replay multiple wiki pages in a manner that makes it easier to identify patterns of representational work that emerge within a group.

Revision history

When the WikiPlayer loads the history of a wiki it reorganizes the information into a multidimensional array to reduce the complexity of the data, simplifying the replay procedure. The data structure describing a revision history

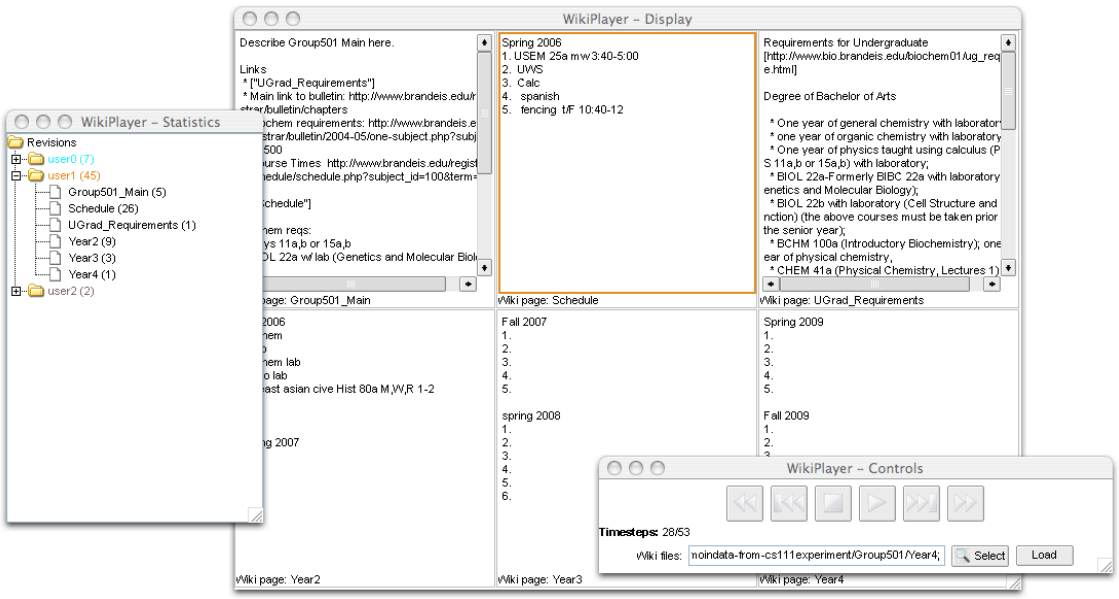

Figure 2. The WikiPlayer user interface

can be explained as follows: A user can make a revision r to a wiki page p at some timestep t. Thus, revisions are denoted by r_{tp} where $R=[r_{tp}]$ is the entire revision history. The size of the revision array is determined by the total number of pages and timesteps, denoted by $t \times p$.

User interface

The WikiPlayer was written in Java and thus can operate on all platforms that can run the Java Runtime Environment (JRE). The application is divided into three independent components [Figure 2]:

- *WikiPlayerDisplay* – Displays/visualizes the wiki revision history.
- *WikiPlayerConsole* – Makes available controls to interact with the history (e.g. play, stop, forward).
- *WikiPlayerStatistics* – Summarizes and displays statistics extracted from the history.

WikiPlayerConsole is the window that contains menu options to load the history into the player as well as "DVD player style" controls that can be used to move around the revision history just like one would be playing his or her favorite movie in a DVD player. Once the history has been loaded into the WikiPlayer it is possible to play it forward/backwards, stop, step forward or backwards, or even fast forward/backwards. The goal is to make the interaction with the revision history as simple and intuitive as possible.

The design of the WikiPlayer enables compatibility with all types of wikis. The *WikiPlayerConsole* asks the *LoaderManager* to parse the revision history on its behalf. The manager uses auxiliary "loaders" (Java classes) written specifically to parse a history from a certain wiki type. Following a special architecture required by implementing special interfaces each loader returns the history in a structured format, uniform for all wikis, that the manager understands and can use to prepare the history data for playback. The preparation involves e.g. synchronizing histories of individual pages and assigning each user with a pseudonym and a unique color. Currently the player supports MoinMoin wiki and MediaWiki (the wiki used on Wikipedia).

The *WikiPlayerStatistics* component receives the prepared history from the manager. It examines the entire revision history and generates useful statistics such as the total number of edits per page per each individual user. The significance of each contribution can be measured by comparing the size/length of each revision to the overall content length of the wiki page. This enables analysts to distinguish between users that made a lot of edits with little content and users that made few or many edits but in either case contributed significantly to the content of the wiki. The statistics are distributed onto each individual user and presented using their pseudonym and color in the *WikiPlayerStatistics* window.

Once a revision history has been loaded the controller notifies the *WikiPlayerDisplay* window so it can initialize itself. The display is divided into subsections similar to a checkerboard each square representing a wiki page. Different groups construct different numbers of wiki pages, so the size of the checkerboard for a given group will vary accordingly. Upon starting the replay each contribution made by a user to a page is easily tracked by highlighting his or her changes to a page using the assigned color. This facilitates identifying the contribution and keeping track of his work on the wiki project when he moves about the different pages as the group collaborates. During playback the revision made to a document is visualized on its respective document item on the display window and all

other documents that were not modified at this timestep retain their appearance from the previous timestep.

EXPERIMENT: DATA ON ONLINE COLLABORATION

In the fall of 2005, a course in Computational Cognitive Science was taught at the Department of Computer Science at Brandeis University. There were 28 students in the class, a mix of graduates and undergraduates; the population included non-majors. The course focused on introducing students to theoretical material on online cooperation and joint sensemaking [1].

The class was divided into teams of 2 to 5 students. The experimental task was organized in such a way that it would encourage online sensemaking and collaboration. The experimental platform used was CEDAR, an application wrapper around a wiki web site that provides a few additional collaborative tools and enables replay of the collaborative wiki assignment at the user interface level [l]. The goal was to construct a collaborative task where there were multiple dependencies among the various subtasks. Each team acted as an academic advisor that was required to develop at least a two-year schedule of classes for an incoming freshman. Part of the task was to post the student's schedule for each year in chronological order within a certain amount of time.

The groups collected information from sources on the World Wide Web including course listings and requirements from the University Registrars office. They combined this information with their own experiences and preferences developed through the collaborative activity and constructed the class schedule in the form of a Wiki page, using MoinMoin wiki. Halfway through the session, the teams were interrupted and informed that their freshman advisee had decided to add a specific minor to his or her schedule. When they were finished, the teams were required to submit the schedule to "the student". The student critiqued the schedule, and always found some changes that needed to be made – usually this meant a change in course selection or sequence of courses – forcing the team to rework the schedule. This required the teams to collaboratively revisit older parts of the schedule and modify them while continuing to work on planning the unfinished semesters.

REPRESENTATIONAL WORK

CEDAR was developed using homegrown groupware engineering toolkits THYME and Sage [5]. CEDAR automatically produces replayable transcripts of the online practice. A VCR-like replay device, provided by CEDAR, allows the analyst to review all the online activity of the subjects at the user interface level, searching for domain meaningful events like a chat event or a web browser event. More detail on the development and usage of CEDAR are provided in [7]. One disadvantage of the replay device in CEDAR is that reviewing the representational work that the students do on the wikis is cumbersome.

We used the WikiPlayer to replay the revision history of the wiki that the students used in the academic advisor assignment; in total we replayed the wiki history of 11 student teams. Our observations revealed some fascinating representational work patterns employed throughout the collaborative activity. We investigated four interesting representational activities in more detail. These patterns represent the techniques that the groups developed during the collaboration to deal with the hindrances caused by the nature of the task, information used, and (lack of) support for representational work. The teams often collaborated on representational activities to deal with interruptions to the work process like the introduction of a minor for the freshman. Each pattern of representational work can be characterized by a distribution and division of labor, the management of collaborative task, and the organization of representational content. We will now describe in more detail the main representational activities that we observed (see also Table 1).

Representational information transfer (mapping)

To successfully finish the class schedule the teams had to locate information on the web from several sources, including the University Registrars' Office and the web pages for various departments. An effective strategy was to transfer information from a web page to a wiki page, while perhaps adding some additional organizational structure. This strategy was effective because it made the information more readily available for continued access.

Representational task management

In some cases, we observed groups creating checklists on a wiki page. For example, the checklists were used to prevent course duplications and to manage group effort. The subjects sometimes used information copied from web pages on to wiki pages as a checklist. Thus, we see the propagation of representation as it is transformed and modified to meet the demands of alternate tasks: first as a statement of requirement, then organized to make repeated access more efficient, and finally as a "checklist" that would insure the completion of all elements of the task.

Students annotated the checklists using words such as "DONE", "FINISHED", or marked entries with asterisks to keep track of the activity. On some occasions they even used links to the wiki pages as markings indicating that a certain requirement had been fulfilled in the schedule. The link would then take the user to the wiki page were the requirements were satisfied.

Representational task organization

Some groups decided to employ a "divide-and-conquer" strategy to finish the assignment. Each group member was assigned a specific task that he or she only worked on during the entire assignment. Up to a point this strategy is

P_i	Pattern title	Description	Grading policy
1	Rep. information transfer	Subjects transfer information from web pages to Wiki	1 pt. assigned for trying 3 pts. for each requirement transferred -1 pt. for each requirement section sharing a page.
2	Rep. Task management	Track progress using checklists	0 pts. for no lists. 5 pts. for lists sharing a wiki page. 10 pts. for detailed list on a separate page.
3	Rep. task organization (collaboration)	Multiple subjects edit each Wiki page.	For each team, the percentage of wiki pages where two or more participants collaboratively edited the page.
4	Rep. system organization (wiki)	Hierarchical structure of Wiki	0 pts. for no hierarchical structure. 5 pts. for using some form of hierarchical structure sharing a wiki page. 10 pts. for using well organized structure on a separate page.

Table 1. Representational work pattern grading policy.

effective. The problem is that there are dependencies, for example, between requirements, and consequently too much division of labor can potentially produce scheduling conflicts that lead to excessive repair work. Thus, many teams avoided explicit "divide and conquer" strategies, preferring to collaborate in detail on the majority, if not all, of the wiki pages. Examples of this kind of fine grain collaboration over representational work include: copyedits of each other's work, multiple subjects adding information to single wiki pages, and the collaborative reorganization of material that was copied from a web page to a wiki page. This kind of anarchic approach to representational work is consistent with the philosophical approach of sites like Wikipedia.

Whether or not the subjects deploy a "divide-and-conquer" strategy is one important determinant of the overall organization of the representational work within a team of subjects. The kinds of representations the subjects worked with, both number and kinds of wiki pages, is another significant determinant of how the collective representational work is organized. How a group works out the details for each one of these organizational tasks largely defines the flow of representations as the group performs their task.

Organization of the representational system

We used the WikiPlayer to identify different kinds of organizational structures for the collection of wiki pages produced by a single group. The teams mainly used two different structures to organize their wiki pages. Some groups only had a few pages, and few if any links between them resulting in a very flat and "horizontal" structure of independent pages not easy to navigate. Other teams divided information onto many pages (representations) adding a hierarchical organization to the structure of the wiki, making it easier to navigate, with relevant information grouped together at the same or close locations on the wiki. The latter organizational scheme often included a main overview page linking to relevant sub pages and/or related pages.

A QUANTITATIVE MEASURE

Our analysis focused on the four representational work patterns described in the previous section.

For each work pattern we assigned points on the scale of 0 to 10. 0 was the lowest amount of effort put into a representational activity by a team and 10 indicated the highest amount of effort. The points were assigned differently to each pattern and the grading strategy is explained in Table 1.

To measure the total amount of representational work for each team we calculated a score, on a scale of 0 to 100, by summing the scores assigned to each representational work pattern that a team performed. We determined that all of the representational work patterns were equally important - each weighing 25% of the total score (denoted by $w=2.5$).

The formula to calculate the total amount of representational work is shown below.

$$\sum_{i=0}^{t} wP_i$$

Where t=4 (total no. of patterns)

For an integer $0 \le i \le t$

To explain the evaluation we will use one group as an example. One successful group transferred all of the requirements earning 10 points for the first pattern, however major and minor requirements shared a single wiki page, which means that two points were subtracted resulting in a total score of 8 for pattern P_1. Using the copied information they created a new separate wiki page for the sole purpose of being used as a checklist to manage the ongoing activity earning them another 10 points for pattern P_2. This group created in total 10 wiki pages and two or more team members collaboratively added content to 6 pages (or 60%) thus the overall score for pattern P_3 is 6 (since all pattern grades are on a scale of 0 to 10). This group put great effort into organizing the wiki into a good hierarchical structure including an overview link page into the structure that earned them 10 points for pattern P_4. Since all the patterns constitute representational work we calculate the final amount of representational work done by the group by summing the scores for the individual patterns. Since the final score is on the scale of 0 to 100 we need to multiply each individual score by the weight factor w that determines how much each score weighs (percentage of the

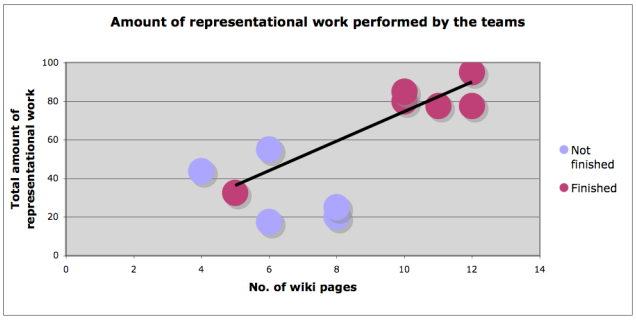

Figure 3. Total amount of representational work.

final 100 points). We calculate the final score according to our formula above:

$$2.5P_1 + 2.5P_2 + 2.5P_3 + 2.5P_4 \quad \text{therefore}$$

$$2.5 * 8 + 2.5 * 10 + 2.5 * 6 + 2.5 * 10 = 85$$

Thus, the total amount of representational work that the group did during the assignment was 85 on the scale of 0 to 100.

RESULTS

Analyzing a revision history for a wiki is helpful to gain insight into how online collaboration is performed. The WikiPlayer proved to be a useful analysis tool to identify collaborative work patterns and how they are influenced by representational activities carried out by the teams within the representational system. We will now present some statistical corroboration to support our observations.

6 of the 11 student teams successfully finished the assignment - producing a complete and detailed class schedule for the incoming freshman. Those teams that finished tended to be more creative at using the wiki in various ways to manage their representational task [Figure 3].

Groups that finished

The 6 teams that finished were more effective at converting both their collaborative task and collaborative meta-task into representational work [Figure 3]. As a result, their representational system was richer and more effective. There was one exception to this rule. The outlier team included only two students, thus the coordination of their effort was less difficult than that of the other groups that finished the task, who had 3 to 5 team members and consequently more complicated coordination problems. In other words, the structure and organization of the flow of representation within a team is, not surprisingly, dependent on the size of the team. As the size of the team varies, the kinds of coordination problems change, and consequently alternate schemes for the representational task become more effective. With only two team members, there are fewer coordination problems, and thus the introduction of representational work to manage the coordination of the activity is not necessarily required.

Figure 4. Wiki page used as a checklist.

All the successful groups did significant amounts of copying from external websites to the wiki. A review of the data using the WikiPlayer showed the successful groups selectively copying only relevant and useful material. These teams also put effort into re-organizing the information once it was copied to a wiki page to better support the use of the information. For example, copied information that was subsequently used for a checklist, was modified by structuring the information to make it easier to annotate with checkmarks.

All of the successful groups except for the outlier relied on using representations to keep track of the advancement of the task, using a wiki page as a checklist. Figure 4 shows an example of how the groups used wiki pages for task management. This group used the wiki page describing general university requirements, that they transformed from an external website, as a checklist with a mark of "x" as an indication that the requirement had been met in the schedule.

The successful teams collaborated on average on at least 50% of all the wiki pages that they created. About half of those teams collaborated on no less than 70% of their wiki page collection. Overall the level of collaboration amongst the successful teams was much higher than the teams that did not finish; this is understandable since two teams out of the five did not collaborate on a single page.

The majority (83%) of the successful teams also spent a lot of their collaborative work on organizing and structuring the representational system including the information mapped from external representations. Dividing up the content and grouping related information together on different locations on the wiki and linking them together forming a hierarchical representational structure proved to be very valuable to the teams as this enabled team members to quickly locate relevant information.

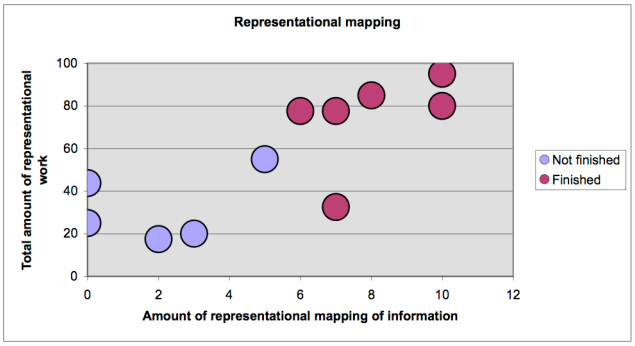

Figure 5. Total amount of representational mapping of information.

Ineffective Groups

One consistent failing of the ineffective groups is that they did not use the wiki as a caching device for collecting information from web pages. 50% of the unsuccessful teams did no transfer of information (mapping) between representations whether it was between internal wiki pages or from external sources to the wiki as shown in Figure 5. Those that tried did very minimal copying often transferring non-useful information or not enough.

Further review of the data using the WikiPlayer also shows that that the unsuccessful teams did not create checklists to manage their collective work. Groups without checklists often had individuals repeat work that had already be done by another team member.

Furthermore, by reviewing the data through the WikiPlayer we observed that a "divide-and-conquer" strategy was not optimal. It did reduce the representational work for the group, but at a high cost. Because of the many dependencies that exist between required courses and the scheduling conflicts between courses, the "divide-and-conquer" strategy produced numerous scheduling problems that required enormous amounts of work to fix.

About half of the unsuccessful teams did not organize the wiki pages into a hierarchical structure. This was also a suboptimal strategy. When the ineffective groups introduced hierarchical structure it was not to organize the wiki pages, rather it was embedded into the content of some wiki page, thus reducing the impact on the flow of information within the representational system.

Hierarchical Clustering

We also used hierarchical clustering methods to group the data together and visualize trends in the collaboration. We wanted to see if other analysis tools would corroborate our observations of the representational work patterns.

We used the Hierarchical Clustering Explorer (HCE) developed at the Human-Computer Interaction Lab at the University of Maryland [9]. This tool can be used to analyze multidimensional data sets such as the statistics we extracted from the wiki revision history using the WikiPlayer.

The software applies the hierarchical clustering algorithm without a predetermined number of clusters, and then enables users to determine the natural grouping with interactive visual feedback. To use the tool we transcribed by hand the statistics obtained from the WikiPlayer and from a brief "by-hand" analysis onto a tab-delimited format that could be loaded into the tool. HCE groups teams together based on their score on different evaluation parameters.

Using the Hierarchical Clustering Explorer (HCE) we were able to identify similarities between the teams based to the evaluation parameters they received from our analysis on representational work. The parameters included the following:

1. Did the group finish the assignment?
2. Team collaboration on wiki pages.
3. Maximum representational mapping by the team.
4. Minimum representational mapping by the team.
5. Hierarchical structure of the representational system.
6. Was the class schedule divided onto semesters per page?
7. Was the class schedule divided onto years per page?

As we expected, the results from HCE supported our observations. HCE presented the results graphically where all the teams were automatically divided into two separate groups, one that included teams that finished the assignment and one that included unsuccessful teams. The clustering shows that there is a clear connection between the success of the assignment and how the teams organized their representations, the propagation of representations between alternate forms, and the division of representational work. The teams that were able to convert most of the collaborative activity into representational work all finished the assignment.

FUTURE WORK

We believe that the WikiPlayer can evolve into a powerful analysis tool for online collaborative activities mediated by a wiki web site. The observations drawn from research using the WikiPlayer also apply to other kinds of collaborative groupware. The WikiPlayer can be used as a tool for teaching transcript analysis to students in classes on Human Computer Interaction (HCI) or Computer Supported Cooperative Work (CSCW). Students can also use the WikiPlayer in order to improve the management of wiki-mediated collaborative tasks.

In its current form the tool requires some "by-hand" analysis to further enrich the results that can be automatically generated and observed through the WikiPlayer. We plan to automate many of the analyses we currently do by hand and explore alternate methods for visualizing the flow of representations. We also plan to include other features of the wikis - such as the "talk pages" of MediaWiki (Wikipedia) - into the user interface of the WikiPlayer. Including the commentary between the collaborators and aligning it with a display of the rest of the

representational activity within the group is likely to reveal other interesting patterns of representational work.

CONCLUSIONS

In this paper we discussed our approach to evaluating online collaborative work in terms of the representational system. The WikiPlayer simultaneously replays the entire state of the representational system at the same time by aligning the individual revision histories for each page revealing the shape of the entire representational system at each timestep and how it evolves through time as a result of the representational activities.

Using the WikiPlayer we were able to replay the entire revision history of a collaborative wiki project and identify four different representational work patterns. The way the WikiPlayer replays and visualizes the revision history of the representational system allowed us to closely examine how the groups carried out these representational activities. Without the WikiPlayer the examination of the large amounts of data we had collected would have been a truly daunting task. The results of our analysis showed a significant relationship between how each team organized their representational system and work, and how effective they were at achieving their collaborative aim.

Building analysis tools such as the WikiPlayer allows groupware researchers and designers to dig deep into the collaborative activities and study how the collaboration is carried out. The "over-the-shoulder" view of the collaboration gives a unique perspective of the online cooperative work that presents a new range of detailed information about the organization and execution of the collaborative work, which is otherwise hard or impossible to obtain. The capability to better model the mechanics of online collaboration is a precursor to a more effective approach to designing online representational systems.

The WikiPlayer also has great value in the educational arena. Technology like wikis can help students to collaborate on projects and assignments. Thus, the WikiPlayer can be used by students to monitor the progress of their collaborative work, and by teacher as a method of assessment. The WikiPlayer also has great practical value for teaching HCI students how to analyze transcripts.

ACKNOWLEDGMENTS

Thanks go to Craig Fuchs for comments and feedback on the design of the WikiPlayer. Special thanks goes to the fall 2005 class of Cosi 111a for providing data for this research.

REFERENCES

1. Alterman, R. The Impact of "Reflective HCI" on teaching in the Classroom. To appear in *CHI 2006 Workshop on "Reflective HCI"*. April 2006.
2. Clark, H. H., *Using language*. Cambridge: Cambridge University Press, 1996.
3. Hutchins, E. (1995a) How a cockpit remembers its speed. *Cognitive Science* 19(3) 265 – 288.
4. Hutchins, E. (1995b) *Cognition in the wild*. Cambridge, MA: MIT Press.
5. Landsman, S. A Software Lifecycle for Building Groupware Applications: Building Groupware On THYME. PhD thesis, Brandeis University, 2005.
6. Larusson, J.A., Tulinius, V.O., & Alterman, R. Measuring the representational work of pilots during an approach briefing. *Technical Report CS-05-261.* Available at http://www.cs.brandeis.edu/~johann (2005).
7. Larusson, J.A., Alterman, R. Integrating collaborative technology into the interdisciplinary classroom. *Technical Report CS-06-262.* Available at http://www.cs.brandeis.edu/~johann (2006).
8. Leuf, B., Cunningham, W. The Wiki Way. Adison-Wesley, 2001.
9. Seo, J. Information Visualization Design for Multidimensional Data: Integrating the Rank-By-Feature Framework with Hierarchical Clustering. PhD thesis, University of Maryland, 2005.
10. Viegas, F. B., Wattenberg, M., & Dave, K. Studying Cooperation and Conflict between Authors with History Flow Visualizations. In *Proc. CHI 2004,* ACM Press (2004), 575-582.
11. Article on Wikipedia: http://en.wikipedia.org/wiki/Wiki
12. MoinMoin wiki site: http://moinmoin.wikiwikiweb.de/
13. Wiki Wiki Web http://c2.com/cgi/wiki?WikiWikiWeb.

Larger versions of pictures in the paper