REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
07-04-2006

2. REPORT TYPE
Final Technical Report

3. DATES COVERED *(From - To)*
Nov. 2001-March 2005; April 2006

4. TITLE AND SUBTITLE
Coordinating Representations

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-02-1-0131

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Richard Alterman

5d. PROJECT NUMBER
05PR04444-00

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brandeis University, Computer Science Dept.
415 South Street
Waltham, MA 02454

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Brandeis University
Office of Sponsored Programs
415 South Street
Waltham, MA 02454

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release: Distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project proposed a cognitive framework that explains how participants in a cooperative activity coordinate and jointly make sense of an unfolding situation. The principles that are derived from this framework guide the construction of methods, tools, and interface technology that enable the cognitive engineering of online environments that structure the activities of distributed planners and actors in a manner that reduces the work involved in sharing meanings, understandings, and assessments.
A toolkit has been developed that enables the construction of groupware environments that automatically produce complete transcripts of online user behavior that are replayable by a VCR like device. Interaction analysis methods have been created that can be used to assess problem areas in an online collaboration. Coordinating representations are then developed that enable the team to more effectively share a common assessment of their cooperative task. We have begun to investigate methods that leverage the representational content of coordinating representations in order to provide some adaptive capabilities to the groupware system.

15. SUBJECT TERMS
Coordinating repreesentation, Technology-mediated collaboration, interaction analysis, Transcript and Replay

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Richard Alterman
U	U	U	UU		19b. TELEPHONE NUMBER *(Include area code)* 781-736-2703

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

20060414065

Contract #
N00014-02-1-0131

Modeling Computer-Mediated Joint Activity

PI: Rick Alterman, Computer Science Department, Volen Center for Complex Systems, Brandeis University, Waltham, MA 02454. alterman@cs.brandeis.edu

Executive Summary

This project proposed a cognitive framework that explains how participants in a cooperative activity coordinate and jointly make sense of an unfolding situation. The principles that are derived from this framework guide the construction of methods, tools, and interface technology that enable the cognitive engineering of online environments that structure the activities of distributed planners and actors in a manner that reduces the work involved in sharing meanings, understandings, and assessments. Validation for this work has employed the methods of Cognitive Science and Computer Science. One set of experiments confirm the utility of the online environments that are created. A second set of experiments measures the cost and effectiveness of the cognitive engineering techniques.

A toolkit has been developed that enables the construction of groupware environments that automatically produce complete transcripts of online user behavior that are replayable by a VCR like device or analyzable using quantitative techniques. The toolkit can be used to rabidly construct online environments that are tailored to the task requirements of a team of collaborators.

Interaction analysis methods have been created that can be used to assess problem areas in an online collaboration. Coordinating representations are then developed that enable the team to more effectively share a common assessment of their cooperative task. These methods have been experimentally tested. In the classroom, we have demonstrated that the analysis methods are teachable. A set of visualization tools have also be constructed that enable the analyst to view the online interaction using an assortment of representations.

We have begun to investigate methods that leverage the representational content of coordinating representations in order to provide some adaptive capabilities to the groupware system. An experimental study demonstrates that coordinating representations can provide a means for some automatic intent recognition. The experiment confirms that users will use the adaptive component that is created and that the use of that component improves their performance.

Project Summary

Objectives

Develop cognitive theory of distributed collaboration among a heterogeneous team of actors. Theory explains how collaborators share a common understanding of their cooperative enterprise despite the dynamics of the situation. Framework provides principles and methods for cognitively engineering online collaborative environments that support military planners in the rapid construction of a response to a time-critical situation.

Develop principles, methods, and tools for cognitively engineering computer-mediated collaborative activities that support the joint sense-making and assessment activities within a team of collaborators.

Use replayable computer-mediated activities as a method and basis for studying and cognitively modeling collaboration.

Approach

This project developed a cognitive framework that explains how participants in a cooperative activity coordinate and jointly make sense of an unfolding situation. The principles that are derived from this framework guide the construction of methods, tools, and interface technology that enable the cognitive engineering of online environments that structure the activities of distributed planners and actors in a manner that reduces the work involved in sharing meanings, understandings, and assessments.

Representation, interaction, and mediation are the critical elements of the cognitive theory of team collaboration. A key idea is that the sense a team of collaborators makes of their engagement depends on the (coordinating) representations that mediate their interaction. How the team construes the situation, their common understanding, the ways in which they think and reason about their collaboration, directly depends on the set of pre-designed representations that mediate their cooperation.

The principles and methods we developed to cognitively engineer a computer-mediated activity depend on an analysis of the representational system and the introduction of coordinating representations.

The representational system provides the participants with a set of choices as to how to distribute effort in maintaining a common viewpoint. The representational system *mediates* the cooperation. A design problem is to arrange the representational system to best support the expected recurrent interactions among the actors.

Goals

Develop a cognitive model of how participants in a computer-mediated cooperative activity coordinate, plan, and jointly make sense of the unfolding situation with and without coordinating representations.

Develop a method to introduce structure (a coordinating representation) that organizes the interaction among distributed planners and actors in a way that simplifies ongoing coordination and reduces work.

Experiments, data collection, findings

Cognitive Theory of Intersubjectivity

Intersubjectivity is the basis for all cooperative action. What the participants share, their common "sense" of the world, creates a foundation, a framing, an orientation, that enable collaborators to see and act in coordination with one another.

Theories of intersubjectivity draw on a mix of three basic elements to construct an explanation: biology, representation, and interaction. Suppose there are multiple actors engaged in a cooperative activity. Because of the visual apparatus of each of the actors, the participants "see" the situation in a similar manner (biology). Because the participants behavior is mediated by their prior knowledge of these kinds of situations, their expectations for how the cooperation will unfold tend to correlate (internal representation). The artifacts that are available at the scene of the activity mediate their behavior (external representation). The organization of exchanges provides opportunities for the participants to display their understanding of the situation and recognize and repair breakdowns (interaction).

The dynamic nature of the participants' task to understand each other is a complicating factor for any explanation of intersubjectivity. Each occasion of cooperation is different due to the idiosyncratic and historical character of that particular interaction: the physical environment is different, the actors are different, the prior experiences of the actors are different, et cetera. At a second level, over extended periods of time, the design of particular task environments are also changing. Because of the dynamics of the situation, an explanation of the relationship between cognition and intersubjectivity must account for the difference between, the change among, and the "sameness" of, similar types of encounters. Because of the dynamics of the situation, data is required that documents, in detail, a sequence of related interactions, within and across episodes of cooperation, where continuity and change can be observed.

As a part of this research project I developed a model of the reciprocal and dynamic relationship between cognition and intersubjectivity. Below I describe two pieces of the framework. See Alterman (2004) for further details.

Common Ground versus Mutual Ground

Common Ground

Suppose two actors A and B are engaged in a cooperative task. Both actors are grounding, assigning a meaning to, objects and aspects of their common situation. Each participant is also familiar with conventions for acting and cooperating. They also have expectations about the types of participants and their typical behaviors. As they proceed, the participants co-construct a shared sense of their joint enterprise.

One way to model the common understanding the emerges between actors as they collaborate is in terms *common ground.* Typically common ground is formally defined in terms of mutual belief about some proposition *P*. There are various formulations of this idea, but each of them depends on each actor mutual believing in some proposition, mentally represented by each interactant in an identical or overlapping manner.

The criteria of identical and/or overlapping mental representations is a very hard criteria to imagine working in most (if not all situations). Suppose A is a leader and B is a novice. Or A has expertise in one area and B in another. Suppose both actors are multi-tasking. What is the likelihood that they have exactly the same identical sense of their common situation? Is this strict criteria necessary to effectively and efficiently accomplish joint projects?

In my earlier work we demonstrated that effective collaboration, in principle, do not require overlapping or intersecting mental representations (Alterman & Garland, 2001). We constructed a computational model of a team of collaborators learning to work together. Because our study was a computational one, we could examine both the internal memory of each actor and his individual ground. Neither the private representation of prior cooperative behaviors, nor the plans they constructed at runtime, significantly overlapped in their internal representation. Nevertheless, the actors exhibited significant improvements in their cooperative behavior: over time, their cooperative behavior was more effective and efficient, it took less communicative work to achieve their collective goals, the time it took to construct a plan was also reduced, et cetera.

In general, it does not seem necessary that A and B mutually believe they have grounded the exact same *p*. A will believe P_1 and B will believe P_2, and P_1 and P_2 are at best some epsilon away from each other; and at the next point in the interaction, A will believe P_3 and B will believe P_4, and P_3 and P_4 are at best some epsilon away from each other. On the other hand, A and B must have some common sense of their joint enterprise.

The fact that A and B expect similar things about the structure of a recurrent runtime activity can only be true in a very gross sense. It is not necessary that the details of an individual's mental representation are identical nor is it necessary that they overlap in any predictable manner independent of the occurring activity. The *functioning* of the mental representations at runtime, and not the mental representations themselves, is the appropriate basis for modeling the role of mental representations in production/accumulation of intersubjectivity.

Mutual Ground

If the participants in a cooperative activity are co-constructing an intersubjective space as they cooperate, how can progress be made in the interaction if they have not achieved mutual belief in some mental content p?

Two actors participate in a cooperative activity. Both actors are engaged in grounding. Individually they monitor the action. Displays and presentations from one actor to another and external events that are of mutual interest are *mutually grounded* by the participants. The failure of either participant to adequately ground an event, external sign, display or presentation that is believed to be relevant by one or another participant to the situation-at-hand may result in one of the actors invoking a meta-process to interact with the other actor(s) so as to fix grounding problems.

So, A and B are acting. A's behavior may be mediated by some internal frame F_i. If B's actions do not fit into the frame that mediates A's behavior either a new frame is selected by A to internally mediate his behavior, or a meta-process to align private representations of shared activities is invoked. If A has a frame F_i that normally achieves his goal and is consistent/grounded with the actions of B up to that point in the current segment of cooperative activity with B, then A believes that he can use F_i to continue the interaction. Suppose A and B have internal frames that are not aligned and a breakdown occurs, i.e.,

Either

- A's internal mediator F_i cannot ground B's behavior in the frame that achieves A's goal and explains B's behavior.

Or

- B's internal mediator F_j cannot ground A's behavior in the frame that achieves B's goal and explains B's behavior.

A communicative interaction occurs; this is a *meta-interaction*. Since A and B can never directly compare their internal representations of the situation, this meta-interaction is essentially a pointing game. One actor makes a presentation and the other actor either accepts the presentation or indicates that further clarification is needed.

1. A makes a presentation that is intended to re-align the internal mediators for A and B.
    a. The presentation points to what A believes is a commonly known organization of behavior that achieves A's goal and that A believes B will agree to participate in.

2. B accepts A's frame if B can find an internal mediator F_j that achieves B's goal and grounds A's behavior.

There are several alternate methods available for making the presentation. There are also a number of conditions that can be loaded onto this definition, concerning the conditions under which, for example, B finds A's proposal acceptable.

The criteria of mutual ground is strictly less than the criteria of common ground. Every case of common grounding is a case of mutual grounding, but vice versa is not the case.

With the criteria of mutual ground there is no problem with an infinite regress in the mutual beliefs of the actors: A and B internally mediate their behavior with frames that are consistent with one another. After the fact, A and B may agree that they mutually believed in some p, but that interaction is mediated by a different set of internal frames than the ones that mediate the behaviors of A and B during their cooperation. With the mutual grounding criteria, the conditions under which progress is stalled and a meta interaction occurs is more precisely defined: either A or B has failed to find a frame to mediate their individual behavior.

In an extended activity, a feature of the situation can be grounded each participant in compatible - but not necessarily identical - ways. Or a feature of the situation is grounded by both participants in compatible ways at time t_1 but at sometime later, t_2, when the feature is again relevant, its significance must be re-grounded. With the mutual grounding criteria, grounding is dependent on an internal mediator. At time t_1 and time t_2 the internal mediators of behavior for one, or another, or for both actors may have changed and consequently a breakdown at time t_2 is explainable within the mutual grounding framework.

When one actor lies to another, mutual ground is achieved, not common ground. If one participant is bored or not interested in engaging in an argument, mutual ground accounts for the progress of the interaction, not common ground. Explaining the differences between how two different actors recall a prior interaction is more easily explained in terms of the criteria for mutual ground.

The formulation of intersubjectivity as common ground depends on the participants, given some identical basis b, mutually believing the identical predicate p. Mutual grounding does not depend on the equivalence of either p or b. Given the situation-at-hand, what the participants individually believe need only be functionally equivalent as they proceed with their cooperation.

Mediation and Accumulation

As the size of the intersubjective space that the participants work in increases, there is a reduction in the amount of work it takes to achieve their collaborative and individual aims. One source of growth is that over time the structure of the activity is debugged. A second source of growth is that the functional distance between individual representations of the recurrent activity decreases. Even for routine activities there are always points in the cooperation where the participants must communicate with one another in order effectively continue the activity.

The development of external representations that mediate the interactive work the team does to maintain a common viewpoint is a third source for increase in common understanding within the team. One form of mediation is the creation of conversational structure that organizes the meta-interactions within the team. A second form of mediation is to pre-design the team's task environment, including mediating representations that facilitate the team's efforts to co-construct a shared understanding and assessment of their common situation. In either case, mediating representations provide the vocabulary the team uses for assessing, communicating, and thinking during their collaboration.

Conversational Structure

Improvements in the efficiency and effectiveness can be achieved by introducing structure to the meta-interactions that are expected to occur during the normal course of recurrent cooperation. Actors will develop secondary structure that organizes a recurrent conversational interaction, making the mutual grounding of participants more efficient and effective.

Our everyday recurrent behaviors include numerous conversational structures that have been invented to organize the co-construction of an intersubjective space in which to operate. Many of these are tailor-made for particular contexts: the courtroom, the psychiatrist's office, at the hospital, or the interaction with the cashier at the supermarket.

The invention of conversational structure that organizes specific expected cases of mutual grounding is the first form of the social accumulation of intersubjective space within a community of actors. Using conversational structure to mediate the co-construction of an intersubjective space in which to operate has the advantage of being highly flexible, requiring few additional capabilities other than the wits of the participants.

Initially, in a response to a breakdown, a meta-interaction occurs that re-aligns the private understandings of the participants. In future situations, where one or another actor anticipates the problem may re-occur, the actor will initiate a conversational interaction to organize the flow of the activity. Over time the actors expect that structure as an organization of their communication at that point of the interaction.

The Task Environment and Coordinating Representations

For many recurrent activities a conversational interaction is not the ideal method for organizing the interaction at runtime: a conversational interaction is inefficient, ineffective, or even not an option. A method for accumulating intersubjective space for recurrent activities in these cases, is to embed into the task environment some preferences for organizing conventional behaviors. This "pre-computes" some of the runtime work of actors (Norman, 1991). It also enables the distribution of work across people: the people who design the task environment (and thereby pre-select a structure for the behavior) can be different from the actors who perform the behavior.

The shift from activities that are organized by mediating structures interactively to recurring activities that have a pre-designed organizational structure as part of the representational system for the task simultaneously expands the intersubjective space in which actors operate and transforms the vocabulary the team uses to make sense of the situation.

As a group begins to collaborate, they settle into a routine. Part of the routinization of their activity is making choices about how to represent (and reason about) various aspects of the task. Recurrent breakdowns in cooperation are one trigger for redesign. Engineering improvements to the task environment that facilitate cooperation should result in a reduction and/or redistribution of representational work in maintaining a common viewpoint.

We will refer to an external representation available at the scene of an activity prior to the current activity that was designed to "solve" a specific recurrent problem of alignment as a *coordinating representation* (c.f, Suchman & Trigg, 1991; Hutchins, 1995ab; Goodwin & Goodwin 1996; Schmidt & Simone, 1996). Coordinating representations are intended to mediate the alignment of private representations during the course of a recurrent activity.

The continued use of a particular coordinating representation changes how the actors jointly construe their engagement. The departure/arrival monitor at the airport mediates the "pointing" between airport personnel and passengers. It mediates the alignment of the private assessments of the participants with regards to the procedure and the mutual ground of the activity of their cooperation. Without the departure/arrival monitor the sense the actors make of the situation, the way they reason about the situation, the structure of their activity is entirely different.

Over extended periods of time, when a community of actors mediate their behaviors using a set of coordinating representations, expectations for how to proceed in the situation, a common "sense" of how to orient oneself, an ontology for that kind of situation will emerge. As new actors, and new generations of actors join the community, the view of activity and cooperation implicit in the coordinating representation that mediates the set of common behaviors is further distributed within a community. As actors continue to operate, the set of conceptions about "normal behavior" are leveraged to expedite the construction of intersubjective space at runtime.

Cognitive Engineering of Computer-Mediated Collaborative Tasks

In the military, a common occurrence is for a team of actors to be multi-tasking and collaborating both synchronously and asynchronously. The members of the team may also be in different physical locations, and consequently much of their collaborative work occurs online. Nevertheless they must share information and must be able to establish a common sense of their status of their joint project. These characteristics of the military task are shared by many other kinds of collaborations. An obvious civilian application is disaster relief in the event of, for example, an earthquake. A abundance of more mundane examples with these characteristics can be found in business, engineering, and education.

Given the theoretical framework discussed above, the practical portion of this project was to develop a set of principles, methods, and tools for cognitively engineering computer-mediated collaborations. The key idea is that the sense a team of collaborators makes of their engagement depends on the (coordinating) representations that mediate their interaction. How they construe the situation, their common understanding, directly depends on the set of pre-designed representations that mediate their cooperation. Thus, as a cognitive engineer, the core issue is: how does one go about developing a pre-designed online task environment, a set of mediating representations, that will enable a team collabotors to rapidly achieve a common sense of a rapidly developing situation?

The methods and tools we have created key on the construction of coordinating representations, as that the CR's play a critical role in how the team makes sense of the situation. The goal is to be able to cognitively engineer a set of CR's that will, for example, effectively structure the activities of a distributed team of military planners engaged in the task of rapidly formulating a plan for a rescue mission.

The method my group has developed to cognitively engineer a computer-mediated activity depends on an analysis of the representational system and the introduction of coordinating representations.

1. An existing online practice is grounded in the representational system provided by a groupware system.
2. Transcripts are collected of online user behavior.
3. In order to identify weak spots in the representational system, the analysis of transcripts focuses on identifying routine behaviors, coordination activity, maintenance of common ground, errors, and cognitive load.
4. The analysis of practice is used to improve the online exchange and management of information by some combination of:
    a. Modify the representational system. Add new coordinating representations that make it easier for team members to reach a common understand and assessment of a situation.

b. Use the information collected through the mediation of team cooperation to enable adaptive components that support the representational work and collaborative activity of users.

This methodology can be applied to both synchronous and asynchronous team projects.

Below I will discuss in detail three technical parts of the project:

Transcript & Replay. Our method cognitively engineering computer-mediated cooperation requires a cycle of development. We have developed and tested a toolkit that enables the rapid construction of groupware environments. The key feature of these environments is that a complete transcript of all online activity within the team is captured in a replayable form. Access to this data is critical to developing an analysis of problem areas within a team's assessment and sense-making activities during their cooperation.

Interaction Analysis and CR Design. We have developed and tested two analysis techniques. One technique analyzes the interactive work that the team must do to maintain an common viewpoint of their cooperative task. The second technique measures the cognitive load of team members as they exchange information about the task domain and organize their cooperation. Given an analysis of the team's interaction, it is possible to mend the representational system so as to improve the joint sense-making activities of the team. We have evidence that our analysis methods have wide application, are teachable, and reproducible. We also have evidence that the resulting changes to the representational system are significant improvements.

Adaptive Components to Support Sharing. A critical problem in any online collaboration is to manage the immense quantity of relevant information that the team must rapidly and selectively access during an ongoing situation. If some of the information can be semi-automatically managed by the system, it is a definite plus. In order to add an adaptive component to a computer-mediated environment it is to be able to determine at any given point in time what the team is trying to do. Because coordinating representations mediate the team's joint sense-making, they provide critical information about the intent of the group. We have begun to develop some AI techniques that leverage the representational structure of the CR's in order to facilitate the recognition of user intent. We have evidence that the CR's can be used to support adaptive components that improve the team's performance.

Transcript & Replay

Collecting data that depends on a runtime interaction is not an easy task. Detailed note taking is incomplete, labor intensive to collect, and by its very nature interpretive. Technology has been used to collect interactional data that is more complete and less dependent on the subjective interpretation of the author. In conversational analysis, transcripts of recorded telephone conversations are used as data for analysis (Sacks, Schegloff, and Jefferson, 1974). Video technology has also been used to collect detailed interactional data (Suchman & Trigg, 1991). Both of these kinds of technology achieve greater fidelity in the recording of the interaction.

There are problems, however, with using either kind of these technologies to collect data. Both kinds of technology have very high transcription costs. Recorded telephone conversations would not be sufficient for a study that analyzes how the design of a task environment mediates cooperation. No matter how many video tapes are collected, there may still be relevant activity that is occurring offline. Collecting multiple video tapes alleviates some of this problem but it also introduces a new one: the correlation of multiple tapes is technically complicated and time-consuming. Both of these technologies work best capturing a single episode of interaction. Neither of these technologies can be easily used to conduct a study that strings together several snapshots of cooperative behavior in order to capture the flow, growth, and development of intersubjective space for a set of recurrent activities within a community of actors.

We have developed THYME, a toolkit for constructing groupware systems that automatically produce replay devices and complete replayable transcripts of online user behavior (Landsman & Alterman, 2003). A *complete* transcript is a transcript that encodes the full range of data (from mouse click to plan action). The transcript and replay techniques have been successfully used to collect and analyze data from several different applications.

The THYME framework provides a flexible architecture for building groupware applications from reusable, tailorable, and analyzable components. Reusability of components allow for groupware applications to be built, tested, and deployed quickly. Because building reusable components is encouraged and supported within the framework, groupware developers can add to the collection of reusable components quickly when they build custom components for their applications. Tailorability enables the developer to modify existing components to suit the needs of the application, thereby reducing the cases where a new component needs to be written. Finally, any interaction with a THYME component is transcribed, and thus is replayable for later analysis.

Evaluation

In the classroom

In the Fall of 2002, an upper level Computer Science class in Human-Computer Interaction was taught at Brandeis University. This class consisted of primarily junior and senior level undergraduate students. In this class, teams of three or four students used the THYME framework to implement a groupware application as their term project. These projects were constructed over the course of the semester. During the implementation phase, the class was given access to the complete THYME framework and an example set of applications, including a chat room and a shared whiteboard.

The class was given 35 days to design an application, including interviewing potential users and prototyping the proposed interface. Twenty-eight days were given to implement and test the groupware application. During the final 21 days of the semester, the teams performed usability testing and analyzed the collected transcripts. Based on this analysis, they proposed possible redesigns of their projects.

Twelve of the fourteen teams of students completed applications that were usable and analyzable. All twelve of the working systems produced complete transcripts of interface events and domain actions. Because of the shortness of the semester, we did not give students the opportunity to generate SAGE playback tools. Since that time, we have implemented techniques that enable the automatic generation of SAGE tools.

When the HCI class was taught in the Fall semester of 1999, the THYME framework was not used. In the previous class, the teams were given 49 days to implement their applications, 21 more days than the 2002 class had. Nevertheless, the earlier course resulted in fewer usable applications, with no application as complete as the ones produced by this class. Moreover, the applications that were produced by the 2002 class have been retooled and used for future projects, testifying to the greater quality of these projects.

An experimental platform for another ONR sponsored project

In conjunction with Sara McComb Isenberg School of Management at University of Massachusetts at Amherst, we constructed a groupware application (Workforce) to be used to study group decision making and teamwork. This application needed to be built to allow multi-hour problem solving sessions between three participants, and had to generate a transcript that was analyzable. The application took twelve hours to build, including testing and the creation of a replay application. This application was built by taking an existing application and tailoring it. The CounterStrike Strategy application provided the most similar fit from our existing stable of applications and subapplications. The Workforce application required a chat room and a shared canvas, both of which were already provided by the CounterStrike Strategy application. Modifying the palette of available drawing tools was a localized change in the shared canvas. The application required two sets of custom components. The first was a new shared canvas palette, containing the names of available workers, shown on the right side of the screen. This palette had a different look and feel from the existing canvas palette. Building a new palette with a different interface was done quickly, as all changes were locally to the palette, requiring little integration testing.

Interaction Analysis & CR Design

Given a complete transcript and a replay device, the next task is to analyze the interaction. We have developed two sets of analysis methods. The first method, analyzes the interactional work required to maintain a common viewpoint given a representational system. The second method, measures the cognitive load of team members as they exchange information about shared domain objects or procedures.

Identifying recurrent situations is the first step towards an analysis of the amount of work the team must do to maintain a common viewpoint and assessment. Potentially any recurrent kind of interaction to establish mutual ground within the team can be more effectively and efficiently achieved by re-designing the representational system the mediates the teams cooperation. Trouble spots can be identified by recurring errors, breakdowns, and slow points in cooperation. In some cases, the team will have developed highly structured conversational interactions to work their way through a complicated task of shared understanding.

The second method measures cognitive load by tracking how, when, how often, and for how long, team members talk about shared domain objects and procedures. For example, the average lifetime of a given kind of topic (object or plan) can be determined by an analysis of the referential structure of the discourse. By noting each reference to a particular topic of discussion, it is possible to construct a summary of referential structure that can be used to model the cognitive effort that the construction of, and access to, shared information incurs in each participant.

Given an analysis of the existing practice among a group of users, it is possible to introduce shared representations (i.e., coordinating representations) that reduce the cognitive effort and workload among participants. The coordinating representations need not be very exotic, but the have to be relevant and effective. Given the overwhelming amount of information potentially available to a user as she multitasks and the reality of limited screen real estate, it is clear the user does not want a large number of coordinating representations. The analysis methods we have develop focus the designer on developing coordinating representations that fix problematic areas in the exchange of information and the maintenance of a common assessment of a situation.

Evaluation

Testing a re-designed system

We conducted a study to assess the performance of different teams of subjects before and after changes were made to their representational system. One set of teams used the base system. A second set of teams (the CR groups) used a revised version of the base system with coordinating representations added to with intent of making it easier for team members to share an assessment of a situation.. Each set consisted of three groups of three subjects. The groups consisted of a mix of area professionals, mostly in computer-related industries, and undergraduate students; all were paid a flat fee for the experiment.

Each team was trained together for two hours in use of their system, and then solved problems for approximately ten hours. To alleviate fatigue concerns, the problem-solving sessions for each group was split into three four-hour sessions. Subjects were asked to fill out entrance surveys to obtain population data, and exit surveys where they could give feedback about their experience with the system and the coordination issues arising in their team.

A set of random problems was produced, and subjects were given a succession of problems drawn from this set. Groups did not necessarily see the same problems, nor in the same order, and because of differences in performance, did not complete the same number of problems over their ten hours of problem solving. To account for this, a general measure of the complexity of a particular problem was devised. This metric was used to normalize results.

We examined several measures of performance. The ones that are most relevant to gauging the joint sense-making activities of the team are tied to the interactions among the participants. Changes in the relative amounts of communication, errors, and clock time coarsely reflect how much work it takes for the participants to maintain common ground.

The quantitative results presented in Figure 1 compare the performance of Base and CR groups over the final five hours of play for each team; after five hours the performance of the teams had stabilized. These results are also normalized over the computed complexity of the problems being solved. The most significant effect is the 57% reduction in communication generated. Also highly significant is the 49% reduction in clock time. There was also a reduction in system events (mouse clicks, etc.), down 38%. Overall domain errors (errors in performing domain actions which led to a toxic spill) were reduced by 61%.

	Improvement
Communication	57% ($p<0.01$)
Domain Errors	61% ($p<0.2$)
System Events	38% ($p<0.06$)
Clock time	49% ($p<0.01$)
Rounds of Activity	22% ($p<0.35$)

Figure 1: Improvement of CR groups over Base groups; final 5 hours of play

Testing the analysis methods

With the help of a class of students, we ran two experiments to gather data on how well the methodology could be taught and employed on novel domains. In this section we summarize the successes and failures of this vetting of the experimental method.

In the Fall of 2003, a class composed of twenty-one Master's students and upper-level undergraduates were taught the analysis techniques presented in this paper. They applied these techniques to a set of standardized transcripts, which were used to provide feedback about the method and about how well they had learned the methods. The class was then split into groups of two to four students; each student group created problems for pairs of subjects to solve cooperatively. The groups then ran experiments and analyzed data that they generated using the methods developed in this project. From this analysis they were able to draw conclusions about how to alter the representation systems of their experimental applications. Most groups were able to successfully apply the methods to suggest interesting redesign possibilities for their systems.

The students were initially given a groupware system, GrewpTool, consisting of a shared editor, a textual chat, and a shared web browser. The tool provides a shared work environment for two or more users, including a shared text area with text color-coded by author, a chat window, and shared and private web browsers. Actions taken in the system can be replayed using a built-in VCR-like tool, allowing the application of our analysis methodologies. Students were split into groups of two to four and were asked to design an experiment where a pair of users would employ the GrewpTool to collaboratively solve a problem. Topics ranged from "plan a 5-night vacation to Boston" to "the wedding dinner planner" to "create a web page describing the culture of a nation." The students then recruited three or four pairs of subjects, trained them in use of the system, and generated about 10 total hours of use data. From this set of data the students were asked to select a single transcript and apply the methods presented in this paper to analyze the interaction.

Two conclusions from the study were:

The methods apply to a variety of domains. Students were asked to submit ideas for redesigning the GREWP tool, based on conclusions from their analysis. The students were given three weeks to generate and submit designs for new representations to improve user performance in their particular domain, with the requirement that these new designs be motivated using the analysis techniques discussed in class, including those demonstrated in this paper. All of the groups were able to successfully motivate that redesign using these methods. Every group found recurring patterns of coordination and recurring errors in the interaction and used these observations to justify and shape their redesign. In some groups the students also identified the creation of secondary structure by the users. About half of the student groups were able to further refine these design ideas by pulling inferences from the referential structure analysis of their data by making assumptions based on the iota types they identified. Most of these groups employed the full method, computing and comparing various measures (such as iota lifetimes and density of mentions) derived from their data.

Analysts draw similar conclusions from the same data. The students were also asked to perform a referential structure analysis of four standard transcripts to test their analytic skills. These transcripts were pulled from data of undergraduates engaged in a pairs-programming session. Parts of this study had been discussed in class on several occasions, so while the students had not seen the specific data they were given, they were familiar with the domain. After the analyses were performed, we engaged the class in a discussion of the results and methods from this analysis, which yielded strong positive feedback about the utility of the method. In addition to providing students with unambiguous feedback about their ability to perform the analysis correctly, this exercise allowed us to test the inter-coder reliability of the methods presented here. Each transcript was analyzed by five pairs of students. The resulting analyses were qualitatively similar, though there were minor variations in results from group to group. About half the groups matched the expert analysis. Groups usually found comparable iotas and made similar conclusions, even where their analyses differed in detail. These differences can in the main be attributed to differing skill levels between student groups. The appearance of this agreement is a most encouraging sign of the applicability of the method.

Adaptive Component to Support Sharing

Adaptive systems incorporate an "intelligent" component that attempts to infer the user's needs, plans, and/or goals and generates some automated support in response that makes it easier for the user to accomplish their ends in the domain of activity. Adaptive components can potentially reduce some of the representational work team members must do to shared understanding and assessments and establish mutual ground.

We are developing a methodology for adding adaptive functionality to groupware systems. Our focus is on how the addition of coordinating representations can both support user efforts to stay coordinated and simplify user intent inference at runtime. In using coordinating representations to stay coordinated, users produce a stream of structured information that is highly relevant to intent inference. This in turn supports the introduction of powerful adaptations using standard AI techniques. Our approach is applicable to all computer-mediated collaborative tasks.

Two significant features of any coordinating representations that are thus employed, are that using the coordinating representation is coordination work the user wants to do and work that reduces her effort on her immediate task – this is true because of the method we employed to construct and test the coordinating representation. Another significant feature is that the coordinating representation is embedded in the semantics of the domain, and consequently is readily understandable to the user.

Evaluation

To evaluate the effectiveness of adaptive component, we performed a 40 hour study with four groups of three people. The base system for the experiment included three coordinating representations. The participants were divided into two populations of two groups each. One was tested for 20 hours on the system without the additional adaptive component. The other was tested on the system including the adaptive component.

The adaptive component we constructed made predictions about each of the participants' current plan. Each participant has the option to confirm one of five possible goals. After a goal is selected, the user is given an option to have the system automatically generate a plan for that goal. In cases where the goal involves multiple actors, the other actors are invited to join the plan. If all invited actors do not accept the invitation, a plan is not generated.

The participants were divided into two populations of two teams each. One population was tested on the adapted system, which contained all three CRs but did not have the adaptive component. The other population was tested on the adaptive system, in which the adaptive component replaced one of the CRs.

Our analysis sought to answer the following questions:

- Was the adaptive component used? An adaptation can only be useful if people are willing to use it.
- Did it improve performance? Even if an adaptation is used, it may or may not improve user performance.
- Did it reduce user effort? Even if an adaptive component is used, and improves performance, it does not necessarily reduce user effort.

All groups used adaptive component to generate plans within the system. Our evidence show that system predictions were confirmed, and furthermore, plans were frequently requested, accepted and executed. Roughly 71% of the plans requested by users were accepted, and 83% of the plans accepted were actually executed to completion; that is, every step in the generated plan was submitted to the server. Nearly a quarter of all plan steps submitted to the server (where an average game contains about 380 total submitted individual steps) came from the adaptive component. The adaptive component also reduced the number of errors per minute; the biggest reduction (50%) was in the joint error rate.

To quantify the impact of the adaptive component on the amount of effort required during planning, we examined the mean duration between steps submitted to the server for plans that were generated manually, versus those that were generated automatically.

The amount of time taken by users between submitting subsequent steps of automatically generated plans was significantly less than that for both the groups that did not have the adaptive component (52%), and for manual planning phases of activity for groups that did have the adaptive component(57%). This corroborates the error data above, which indicates that coordination was less difficult for automatically generated plans. We conclude from this result that the adaptive component reduced the workload of the collaborators.

Our validation study demonstrates that the adaptive component provided useful and usable adaptive support to users of the VesselWorld system. The adaptive component was heavily used, reduced coordination errors, and reduced cognitive load. These results demonstrate how a practice-based language, based upon the analysis of coordination work in a groupware system, can be leveraged to produce useful adaptations that do not introduce more work for the user. These results provide support for our overall methodology.

Technology Demonstrations

We have done several technical demonstrations over the period of this contract. Numerous computer-mediated tasks have been developed using our methodology. On a separate contract with NSF, we used our principles and methods to develop a groupware application (GrewpTool) that we have used for extensive experimentation in collaborative learning. One of are testbed's VesselWorld was demonstrated at CSCW 2000. The Workforce application we developed for Sara McComb at Umass Amherst is being used as an experimental platform.

Testbed Integrations

Science Accomplishments

The theory of intersubjectivity explains how collaborators share a common understanding of their cooperative enterprise despite the dynamics of the situation. The use of groupware as a basis for scientific enquiry is a methodological innovation. The concept, tools, and methods for cognitively engineering online environments is also significant.

Products, Deliverables, Prototypes

- Numerous published articles.
- Three Ph. D. theses are almost complete.
- Several prototypes have been constructed using the cognitive engineering methods we have developed.
- THYME & Sage – A toolkit for building groupware systems that produce complete replayable transcripts. Brandeis may patent some of this technology.

Impact/Potential Product Transitions

The cognitive engineering techniques we developed will have wide application and is likely to significant impact on the development of the next generation of groupware applications.

Collecting replayable data from online collaborations is a significant methodological breakthrough. It enables the experimenter to conduct a study that strings together several snapshots of cooperative behavior in order to capture the flow, growth, and development of team collaboration over time.

The THYME and Sage toolkit will have wide application in education and research.

Results Summary

Summary of Accomplishments

TECHNOLOGY DEMONSTRATIONS	SCIENCE ACCOMPLISHMENTS	PRODUCTS, DELIVERABLE, PROTOTYPES	IMPACT, PRODUCT TRANSITIONS
Development of numerous groupware platforms	Cognitive theory of the dynamics of intersubjectivity and cognition	Numerous published articles	Cognitive Engineering methods
GrewpTool	Use of groupware as basis for scientific enquiry is a methodological innovation	Three Ph.D. Theses near completion	Brandeis may patent some of the technology THYME & Sage technology
Workforce application for Sara McComb at NSF	Principles, methods, and tools for cognitively engineering online environments	THYME& Sage Toolkit	Data Collection methods

Summary of evaluation

<table>
<tr><td colspan="3">TRANSCRIPT & REPLAY</td><td colspan="4">INTERACTION ANALYSIS & CR DESIGN</td><td colspan="2">ADAPTIVE COMPONENT TO SUPPORT SHARING</td></tr>
<tr><td colspan="3">Students used toolkit to built groupware applications</td><td colspan="4">Experimentally compared base versus re-design system with CR's</td><td colspan="2">Experimentally tested adaptive component</td></tr>
<tr><td>Only 28 days to implement</td><td>12 of 14 groups implemented running systems</td><td>Significant improvements of previous classes</td><td>CR's decreased domain errors</td><td>CR's decreased clock time</td><td>CR's decreased number of system events</td><td>CR's improved team performance</td><td>Improved user performance</td><td>Decreased user work</td></tr>
<tr><td colspan="3" rowspan="2">Transcript & replay used in numerous classes and groupware development cycles</td><td colspan="4">Tested analysis methods</td><td colspan="2" rowspan="2"></td></tr>
<tr><td colspan="2">Methods apply to a variety of domains</td><td colspan="2">Independent analysts draw similar conclusions from same data</td></tr>
</table>

Publications, Technical Reports

Alteman, R. The Dynamics of Cognition and Intersubjectivity. Submitted to a journal.

Introne, J. & Alterman, R. Practice-based Adaptive Systems. Submitted to a journal.

Feinman, A. & Alterman, R. Modeling Cognitive Load in Groupware. Submitted to a journal.

Introne, J. and Alterman, R. Leveraging a better interface language to simplify adaptation. *Proceedings of International Conference on Intelligent User Interfaces*, 262-264, 2004.

Langton, J., Hickey, T., and Alterman, R. Integrating tools and resources: A case study in building educational groupware for collaborative programming. To appear in *Consortium for Computing Sciences in Colleges New England*, 2004.

Garland, A. and Alterman, R. Autonomous Agents that Learn to Better Coordinate. In *Autonomous Agents and Multi-Agent Systems*, 8(3): 267-301, 2004.

Langton, J., Hickey, T., and Alterman, R. Integrating tools and resources: A case study in building educational groupware for collaborative programming. To appear in Consortium for Computing Sciences in Colleges in New England, 2004.

Landsman, S. & Alterman, R. Analyzing Groupware On THYME Technical Report CS-03-234. Department of Computer Science, 2004.

Alterman, R. & Kirsh, D. (editors) Proceedings of the 25th Annual Meeting of the Cognitive Science Society, 1400 pp. LEA 2003.

Feinman, A. and Alterman, R. Discourse Analysis Techniques for Modeling Group Interaction. *Proceedings of the Ninth International Conference on User Modeling*, 223-237, 2003.

Introne, J. and Alterman, R. Leveraging collaborative effort to infer intent. *Proceedings of the Ninth International Conference on User Modeling*, 133-137, 2003.

Landsman, S., and Alterman, R. Building Groupware On THYME. Technical Report CS-03-234, Department of Computer Science, Brandeis University, 2003.

Alterman, R. Coordination of Talk and Action in Collaborative Systems. *Proceedings of* SSGR 2002.

Hickey, T., Alterman, R., and Langton, J. TA Tool. Proceedings of 33rd ACM Technical Symposium on Computer Science Education, 50-51, 2002.

Alterman, R. Talk and Action. Proceedings of Twenty-Fifth Annual Meeting of the Cognitive Science Society, 2002.

Garland, A. and Alterman, R. Learning Procedural Knowledge to Better Coordinate. *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, 1073-1079, 2001.

Alterman, R., Feinman, A., Landsman, S., and Introne, J. Coordinating Representations in Computer-Mediated Joint Activities. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, 43-48, 2001.

Alterman, R. and Garland, A. Convention in Joint Activity. *Cognitive Science*, 25(4) 611-657, 2001.

References

Alterman, R. and Garland, A. (2001). Convention in Joint Activity. *Cognitive Science*, 25(4), 611-657.

Alterman, R., Zito-Wolf, R., and Carpenter, T. (1998) Pragmatic Action. *Cognitive Science*, 22(1), 53-105.

Alterman, R., Feinman, A., Introne, J., and Landsman, S. Coordinating Representations in Computer-Mediated Joint Activities . Technical Report CS-01-214, Department of Computer Science, Brandeis University, 2001. *Proceedings of 23rd Annual Conference of the Cognitive Science Society* , 2001.

Bartlett, F. C. 1967. *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Clark, H. and Brennan, S. (1991). Grounding in Communication. In Resnick, Levine, and Teasly editors, *Perspectives on Socially Shared Cognition*, 127-149.

Clark, H. (1996). *Using Language.* Cambridge University Press.

Cole, M., & Engeström, Y. (1993) A cultural historic approach to distributed cognition. In G. Salomon (Ed.), *Distributed Cognitions,* 1-46. Cambridge University Press.

Drew, P., & Heritage, J., Editors. (1993) Talk at Work. Cambridge University Press.

Feinman, A. and Alterman, R. (2003) Discourse Analysis Techniques for Modeling Group Interaction. *Proceedings of the Ninth International Conference on User Modeling*, 228-237.

Garfinkel, H. (1967). Studies in ethnomethodology. Polity. Cambridge,, UK: Priority press.

Goodwin, C. and Goodwin, M.H. (1996). Seeing as situated activity: Formulating planes. In Engeström and Middleton (editors), *Cognition an communication in work.* Cambridge University Press, Cambridge, U.K.

Hutchins, E., Hollan, J, and Norman, D. (1986) Direct Manipulation Interfaces. In Norman & Draper (editors), *User Centered System Design: New perspectives on human computer interaction.* Lawrence Erlbaum Associates. Hillsdale, NJ.

Hutchins, E. (1995a) *Cognition in the Wild.* MIT Press. Cambridge, MA.

Hutchins, E. (1995b). How a cockpit remembers its speed. *Cognitive Science,* 19:265-288.

Introne, J. and Alterman, R. Leveraging a better interface language to simplify adaptation. *Proceedings of International Conference on Intelligent User Interfaces,* 262-264, 2004.

Landsman, S. and Alterman, R. Analyzing Usage of Groupware: THYME is on your side. Technical Report CS-02-230, Department of Computer Science, Brandeis University, 2003.

Landsman, S., Alterman, R., Feinman, A, Introne, J. (2001) VesselWorld and ADAPTIVE. Technical Report: CS-01-213, Brandeis University. Demo given at CSCW 2000.

Landsman, S. (2004) Building Groupware on Thyme. PhD thesis. Computer Science Department, Brandeis University.

Lave, J. (1991) Situating Learning in Communities of Practice. In Resnick, L., Levine, J., and Teasley, S. (editors), Perspectives on Socially Shared Cognition, American Psychological Association. 63-82.

Minsky, M. (1975) A framework for representing knowledge. In P.W. Winston (ed.), The Psychology of Computer Vision. 211-277.

Norman, D.A. (1991) Cognitive Artifacts. In Carroll, J. (editor), Designing Interaction: Psychology at the human-computer interface. Cambridge University Press.

Norman, D.A. (1993) *Things That Make Us Smart.* Perseus Books, Cambridge, MA

Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:696-735.

Schegloff, E. (1986) The Routine as Achievement. Human Studies, 9: 111-151.

Schegloff, E. (1992) Repair after Next Turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 9:1295-1345.

Schegloff, E. (1991) Conversation analysis and socially shared cognition. In Perspectives on Socially Shared Cognition, Editors Resnick, L., Levine, J. and Teasley, S. American Psychological Association. 150-171.

Schegloff, E. and Sacks, H. (1973) Opening up closings. *Semiotica*, 8:289-327.

Schmidt, K. and Simone, C. (1996): "Coordination mechanisms: Towards a Conceptual Foundation of CSCW Systems Design", Computer Supported Cooperative Work 5, p. 155-200.

Shneiderman, B. Designing the User Interface 3rd Edition. Addison-Wesley, 1998.

Suchman, L., and Trigg, R. 1991. Understanding practice: video as a medium for reflection and design. In *Design at work: cooperative design of computer systems*, edited by J. Greenbaum and M. Kyng. Hillsdale: Lawrence Erlbaum Associates.

Tomasello, M., Kruger, A. and Ratner, H. (1993) Cultural Learning. Behavioral and Brain Sciences 16: 495:552.

Vygotsky, L. S. (1978). *Mind in Society*. Harvard University Press, Cambridge, MA.