The Dynamics of Cognition and Intersubjectivity

Richard Alterman

Computer Science Department
Center for Complex Systems
Brandeis University

2/18/05

Abstract

What the participants share, their common "sense" of the world, creates a foundation, a framing, an orientation, that enables human actors to see and act in coordination with one another. The methods the participants use to understand each other as they act, also produce, either intentionally or unintentionally, partial understandings, lack of understandings, and misunderstandings, which are also a part of what makes for intersubjectivity. Further complicating matters is the idiosyncratic and historical character of each interaction. Nevertheless, the intersubjective space in which actors operate can become richer and easier to produce with the recurrence of behavior.

The first part of this paper develops a cognitive model of intersubjectivity that accounts for the sameness of, and changes between, two similar types of encounters occurring at different times. The second part of the paper presents a ethnographic study of an online cooperation that documents the interactive, constructive, and productive parts of the model.

Introduction

Intersubjectivity – literally "between subjects" – is arguably the organic structure of human cognition. What the participants share, their common "sense" of the world, creates a foundation, a framing, an orientation, that enables human actors to see and act in coordination with one another. Without intersubjectivity there is no human communication, no accumulation of knowledge within a community across generations, no emergence of complex patterns of social interaction, and so on.

Explanations of intersubjectivity are constructed out of a mix of three basic elements: 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 each participant's 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 participants, their goals and prior experiences, the physical context, and the design of the task environment are, in combination, on each occasion of activity, always unique. Because of the dynamics of the situation, an explanation of the 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.

The participants produce at runtime an intersubjective space in which to proceed with their cooperation. The intersubjective space supports the coordination of behavior. The ways in which the participants have failed to understand or have misunderstood each other's viewpoint is also part of the intersubjective space in which they operate. Or to put it differently, the methods the participants use to come to believe that they mutually believe something, also produce, either intentionally or unintentionally, partial understandings, lack of understandings, and misunderstandings, which are also a part of the intersubjective space.

The goal of this paper is to develop a cognitive model of how two actors, with individual subjective viewpoints, somehow come to understand each other and their cooperation sufficiently to make progress. Introducing dynamics complicates the issue somewhat. A puzzling feature of the dynamics is that the intersubjective space does not accumulate across episodes of cooperation between actors, but the pace at which it is produced, and how effective it is at supporting cooperative action does change. The model that is presented explicates the factors that are predictive of positive change in the runtime dynamic of the participants constructing intersubjective space.

In the second part of the paper, a study is presented that documents the interactive, constructive, and productive parts of the model. The study characterizes the flow and dynamic of intersubjectivity within and across episodes of collaboration among different teams of actors within the same community. A methodological innovation is to use an online same time/different place collaboration as a source of data. Because the collaboration occurs online, it is easier to collect data that includes more than a single snapshot of the subjects working together. Because the collaboration occurs online, access to all that the participants share is readily accessible and reviewable.

Intersubjectivity

The term interaction will be used as shorthand for social interaction. I assume there are significant differences between the interactions of an agent with her environment, the interactions between humans, and the interactions among other kinds of species. Arguably it is the case that notions like internal representation and intersubjectivity are only relevant to an analysis of human-to-human interaction or the usage of humanly constructed artifacts.

There are several distinctions observed among representational terms. Representations can be external to the actor or internal, shared or private. Representations can mediate individual behavior or mediate the interaction between two or more actors. The terms frame (Minsky, 1975) and schema (Bartlett, 1967) will be used interchangeably. Any frame can internally mediate the behavior of an actor. The paper is agnostic with regard to the representational form of internal representations. Some kinds of tacit knowledge relevant to the question of intersubjectivity may not be internally represented. I believe that there is no way to account for long and short-term dynamics in human-to-human cooperation without positing internal representations of some sort.

There are a large catalogue of terms that I will use to describe the engagement. Actors are alternately described as participants, interactants, agents, collaborators, and members of a community. The engagement is referred to as an activity, a runtime behavior, a cooperation, a joint behavior, or some combination thereof. Much of the discussion is about recurrent activities and the emergence of conventions: over time, actors create and learn to expect certain kinds of organizational structure for recurrent activities. For recurrent activities, mediating structures are invented that function to support the expected structure of the participants' cooperation.

The interaction among actors may be embedded in a larger set of interactions. A breakdown at one level of the interaction can produce a conversational interaction that works through the breakdown.

There are a number of terms and expressions I use that emphasize that intersubjective space is uniquely constructed on each occasion of activity. "The participants co-construct an intersubjective space in which to operate" or "the participants construct intersubjective space" or "the participants align their private views of the situation" are some examples.

Intersubjective space is used to coordinate activity between actors. It also includes the difference between the private understandings of actors; that these differences can be intended, unintended, and/or produce breakdowns is a relevant consideration.

The Equivalence of Internal Representations

Consider an example of individual action where an actor approaches a door and attends to the doorknob (see Figure 1). There is a three-part relationship to the situation: an *external object* (the doorknob), the frame that predicts the existence of the doorknob (*common knowledge*), and the *grounding* of the doorknob slot from the doorknob frame. The individual has general knowledge about doors, doorways, and doorknobs, that is commonly held within her community, and it internally represented. On a given occasion of behavior, the actor may ground some of her knowledge by *binding* a slot in one or another frame.

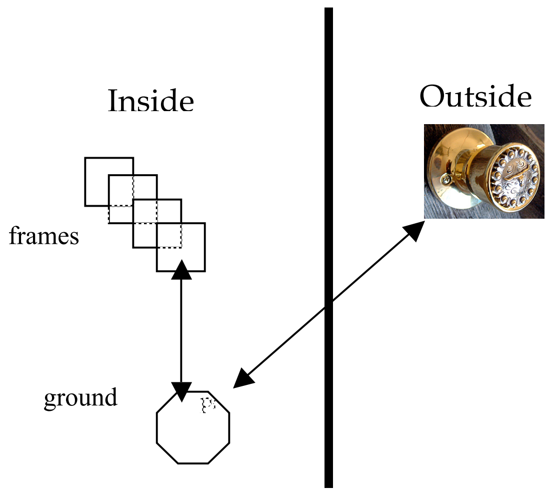

Figure 1: Grounding

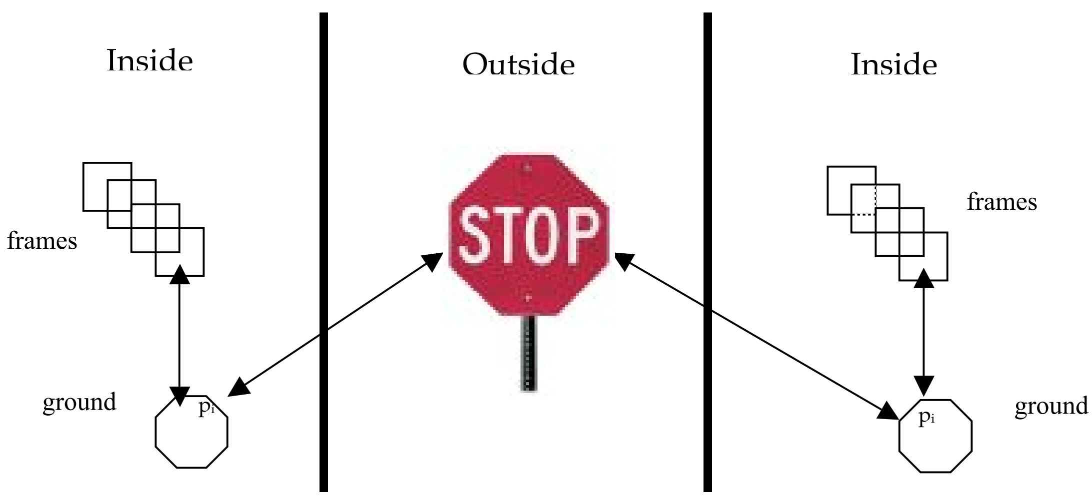

Figure 2: Two Actors Grounding.

This analysis does not state that all parts of an individual's behavior are mediated by mental representations, nor does it state by what process mental representations are grounded, nor how mental representations are represented, only that when grounding occurs, when "sense" is made, this is the sort of thing that is traditionally meant in Cognitive Science.

Suppose there are multiple actors approaching a stop sign at an intersection in the road (see Figure 2). In this case, both actors make sense of the stop sign. There is a physical object at the scene of the activity and both parties attend to it. The physical object has meaning. The meaning that the physical object signifies has bearing on the coordination of their cooperative activity. It mediates the interaction.

In a situation like this, each individual brings to bear a tremendous amount of knowledge. Although each participant, on a standardized test, could identify the traffic laws, the internal representation of each participant is not likely to be a rote memorization of the law. In addition to basic information about the traffic laws on stopping, each participant is also familiar with conventions for acting, under various conditions, when a stop sign is in force at an intersection in the road. Other kinds of relevant knowledge for which each actor has a mental representation concern the types of participants in a traffic situation (other drivers, cyclists, and pedestrians) and expectations about their typical behaviors, rolling stops versus legal stops, rush hour traffic, et cetera. In each case, the meta knowledge that an individual participant employs is contingent on any number of factors -- e.g., her age, skill, confidence as a driver, the performance characteristics of the car – and therefore is unlikely to be identical among drivers.

Only a selection of this knowledge is grounded by each individual during a particular encounter. Road constraints, the heaviness of traffic, and time constraints are all nuances of the situation-at-hand that influence how and what is grounded.

It is unlikely that the "common knowledge" of the participants, which is inextricably tied to their private representations of their individual experiences of driving, is internally represented in an identical, overlapping, or intersecting manner. It is also unlikely, nor is it necessary, that given the nuances of the situation-at-hand, the participants have intersecting or overlapping grounds. One participant, Ludwig, may be in a rush and is grounding his method for piggybacking on the driver in front of him so he can get through the stop sign before it is his turn; and the other driver, Wolfgang, may have noticed a pedestrian with a young child has entered the crosswalk. From the point of view of an outside observer, one could claim that their ground *functioned* as if Ludwig and Wolfgang both grounded the predicate that Ludwig would go first, but if we could examine what they actually grounded, in this particular situation, the sense each of them made of the stop sign is different.

Ludwig and Wolfgang share an intersubjective space sufficient for them to coordinate their activity, but that does not mean they have the same sense of the engagement. There will be ways in which the have different understanding, only some of which emerge during the interaction. In general, intersubjectivity cannot be equated to equivalence of mental representations. There is a wide range of situations in which actors are constructing and sharing an intersubjective space. The analysis of the situation at the stop sign shows that even to achieve coordination between actors it is only necessary that their mental representations of the situation are *functionally equivalent*.

Doers of action

The view of intersubjectivity that developed from ethnomethodology argued against the notion of internal representation as the basis of intersubjectivity. Rather, the focus should be on the organization and flow of social interaction that *produces* a common understanding. " The appropriate image of a common understanding is therefore an operation rather than a common intersection of overlapping sets" (Garfinkel, 1967: p. 30).

From this perspective, an explanation of intersubjectivity should focus on the procedures by which "doers of action" produce shared knowledge (Schegloff, 1993: p. 1299):

> Instead, what seemed programmatically promising was a *procedural* sense of "common" or "shared," a set of practices by which actions and stances could be composed in a fashion which displayed grounding in, and orientation to, "knowledge held in common" – knowledge that might thereby be reconfirmed, modified, expanded and so on. (Schegloff, 1993: p. 1298)

The organization of ordinary conversation provides opportunities for the interactants to display their understanding of the situation-at-hand and also recognize and repair breakdowns of intersubjectivity (Schegloff, 1992:1993). Conversation is sequential; the interactants take turns. In the first position, a speaker presents a contribution to the conversation. In the second position, other participants have an opportunity to display a response. In the third position, the initial speaker can amend her presentation if it did not invoke a preferred response. In this manner, it is the organization of the conversation, the organization of repair in conversation -- the interaction, not representation -- that forms the basis, the framework of analysis, for the intersubjective.

The work of Clark (1996) develops a model of communication that integrates the talk-in-interaction view with the production of *common ground*. In the situation-at-hand, the construction of meaning is a participatory collaboration, requiring both presentation and acceptance phases (Clark & Brennan, 1991). One actor, A, presents an utterance *u*, with the expectation that during the acceptance phase B will provide evidence that B understood *u*. There is a range of evidence that B can contribute that varies in the strength of evidence the contribution provides: continued attention, initiation of the relevant next contribution, acknowledge, demonstration, and display (Clark & Schaefer, 1989). In general, the participants try to minimize the total effort spent on their contribution in both the presentation and acceptance phases of the interaction (*principle of least collaborative effort)*.

During their interaction, the participants are engaged in grounding (Clark & Marshall, 1981; Clark, 1996). The criteria for *common ground* are shown in Figure 3. *P* is a part of common ground if there is a basis *b* that indicates to A and B that *p*. The basis *b* may be that the actors are attending to the same object and they are rational, or it may be their prior common knowledge. This definition includes a caveat that *p* is believed sufficient for the current purposes.

For two people A and B, it is common ground that *p* iff;

1. A and B have information that some basis b holds;
2. b indicates to A and B that A and B have information that b holds;
3. b indicates to A and B that *p.*

Figure 3: Grounding Criteria

In the case of Ludwig and Wolfgang, *p* might be the proposition that "Ludwig is going first", and *b* would include general knowledge about traffic situations. The common ground criteria is met if Ludwig and Wolfgang mutually believe that "Ludwig is going first" before Ludwig crosses the intersection first.

But, as discussed above, Ludwig and Wolfgang can successfully coordinate even if their understandings have not met Clark's criteria of *common ground*. Ludwig and Wolfgang can be aware of each other's activities. Ludwig can believe that "if he piggy backs on the driver in front of him, he can get through the intersection faster" and Wolfgang can believe "I will wait for the pedestrian in the crosswalk to cross the road before I cross the intersection". After the fact, they might agree that they both "thought" before hand that "Ludwig was going first", but that was not necessarily the case before Ludwig crossed the intersection. What the participants individually believed need only be functionally equivalent in order to coordinate as they proceeded with their cooperation. Only after the fact, only in the absence of a breakdown is it known that their intersubjective space was sufficient for the participants to achieve their individual and collective aims.

In general, the methods the participants use to coordinate their behavior also generate, either intentionally or unintentionally, partial understandings, lack of understandings, and misunderstandings, which are also a part of what makes for intersubjectivity. A husband and wife who have a fight and completely disagree who is responsible for what share an intersubjective space, as do a liar and his audience, an expert and a novice, a finance officer and a bank clerk, a quarterback and his receiver, and so on. In each of these cases, the ways in which the participants do not have similar understandings and yet are able to successfully cooperate and coordinate, has as much to do with intersubjectivity as does the ways in which the participants think they have understood each other.

Procedural Framing

Intersubjectivity is located in the interaction, the procedure in which the participants display their orientation towards the collaboration. The organization of the interaction provides the participants opportunities to display, repair, and orient themselves as they proceed with their activity.

At each moment, how the individual acts and/or what they present is guided by the actor's intent or plan. For a cognitive scientist, an actor's plan is a mental representation that mediates the organization of the individual's behavior (Miller, Galanter, and Pribram, 1960). A plan is used to control the order in which a sequence of operations is to be performed: the intended behavior is the unfinished parts of the plan. Any number of behaviors rely on some sort of internal representation to organize the behavior. How much of the behavior can be accounted for by the plan is more complicated. There are several good reasons why a plan cannot account entirely for the actor's behavior: these include the pace of the action, the potential complexity of the representation, and uncertainty (Suchman, 1986; Agre & Chapman, 1990).

The working assumption of this paper is that collaborative and cooperative activity is in part mediated by an internal representation. This will be referred to as the *procedural framing* of the individual actor's behavior: there exists an internal mediator F of each actor's behavior that represents the "sense" an actor makes of the situation from the perspective of her intent.

Atilla and Bambi are acting. Atilla's behavior may be mediated by some internal schema F_i. If Atilla has a frame F_i that normally achieves his goal and it grounds the actions of Bambi up to that point in the current segment of cooperative activity with Bambi, then Atilla believes that he can use F_i to continue the interaction. After the fact, if the cooperation runs smoothly, Attila, if asked, might say he thought there was mutual belief about some factor or another, but that extrapolation may only be post hoc. In many cases, Atilla does not need to actually reason explicitly about Bambi's intentions or beliefs in order to successfully proceed with the cooperation. At runtime, it is sufficient that Bambi's behavior fits into an internal frame that Atilla uses to mediate his behavior and achieve his goals.

If Bambi's actions do not fit into the frame that mediates Atilla's behavior either a new frame is selected by Atilla to internally mediate his behavior, or a communicative interaction is invoked to align private representations of the shared activity.

Suppose Atilla and Bambi have internal frames that are not aligned and a breakdown occurs, i.e., either

- Atilla's internal mediator F_i cannot ground Bambi's behavior in the frame(s) that achieves Atilla's goal and explains Bambi's behavior.

or

- Bambi's internal mediator F_j cannot ground Atilla's behavior in the frame(s) that achieves Bambi's goal and explains Bambi's behavior.

A communicative interaction occurs. Since Atilla and Bambi can never directly compare their internal representations of the situation, the communicative interaction is essentially a pointing game (Garfinkel, 1967). One actor makes a presentation and the other actor either accepts the presentation or indicates that further clarification is needed; this is both Schegloff's analysis and Clark's analysis.

1. Atilla makes a presentation that is intended to re-align the internal mediators for Atilla and Bambi.
    a. The presentation points to what Atilla believes is a commonly known organization of behavior that achieves Atilla's goal and that Atilla believes Bambi will agree to participate in.
2. Bambi accepts Atilla's frame if Bambi can find an internal mediator F_j that achieves Bambi's goal and grounds Atilla's behavior.

The introduction of procedural framing achieves several things. With procedural framing, agreements and coordination depend on only the functional equivalence of mental representations. Differences of perspective, roles, and expertise are also part of the procedural framing of intersubjectivity. When one actor lies to another, if one participant is bored and not paying attention but wants to continue the interaction anyhow, procedural framing accounts for the progress of the interaction. The differences between how two different actors recall a prior interaction can also be explained using procedural framing. Suppose 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, a breakdown occurs, how is that explained? 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 occurred.

Points of Coordination, Dynamics, and Mediation

For a recurrent cooperation, the size and speed at which an intersubjective space is constructed at runtime will change. As a cooperative task recurs, the actors design and debug a procedure for that activity. As the actors settle into a routine the functional distance between their mental representations of the expected structure of the activity will decrease.

In Alterman & Garland (2001), we presented a computational cognitive model of members of a moving company working and acting together (MoversWorld). In MoversWorld, each of the actors reasoned, planned, and learned independently. They interacted with one another by either talking or acting. The behavior of an individual MoversWorld actor was guided by the adaptation of recalled prior episodes of experience. Individually each actor learned through experience, storing in memory a representation of segments of coordinated (and uncoordinated) behavior that were extracted from an "execution trace" of the individual's activity. The segments of behavior were organized, indexed and retrieved by the points in the segment when the agents interacted, the points of coordination.

Conventions of behavior emerged within the MoversWorld community. Over time, each participant began to act in ways that would simplify coordination. The actor began to expect that certain procedural structures for cooperation would work better than others. But the structure of a recurrent behavior was not uniquely determinable independent of the occasion and procession of a given activity. Certain points of coordination tended to be realized on each occasion, but in different manners. One time it emerged one way, a second time a different way.

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 in MoversWorld 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, and so on.

A key point to the MoversWorld model is that with recurrence of a specific kind of cooperation, the actors debug the design for that behavior, and their expectations better align, but there are always *points of coordination*: moments in the interaction where the participants must mark, confirm, or navigate their progress through their private expectations of how the collaboration will unfold.

Because the actor's begin to anticipate certain points of coordination, they can create mediating structure that organizes a recurrent point of coordination so that the interaction will run smoothly at that expected moment of the cooperation if, and when, it emerges.

The study presented in the second half of this paper documents the development and relation between two forms of mediation:

- conversational structure that organizes a recurrent point of coordination
- a re-design of the task environment that introduces a sign to organize a recurrent point of coordination

Initially, in response to a breakdown, a conversational 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 actors 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.

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. However, for many recurrent activities a conversational interaction is not the ideal method for organizing the interaction at runtime. A conversational interaction can be inefficient, ineffective, or even not an option.

An alternate method for making the interaction at a point of coordination more efficient and effective 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.

Debugging and the convergence of expectations about the structure of recurrent activities result in reductions in the work it takes to produce an effective intersubjective space. Both forms of mediation potentially increase the speed and effectiveness of the interaction among participants at the expected points of coordination.

Conversational Structure

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 at a given point of coordination.

For example, there is a core opening sequence to the initial stages of a telephone conversation (Schegloff, 1986); see Figure 4. The phone rings, summoning someone to pick up the phone. He answers by saying something like "hello" or "This is the Computer Science Department". Next the participants identify themselves and/or recognize each other (e.g. "This is Jim", "Hi, Jim"). Then they greet each other, ask how each is doing, and then, normally, the caller introduces the first topic of conversation.

1. The summons/answer sequence.
2. The identification (and/or recognition)
3. Greetings (i.e., ratified mutual participation)
4. 'Howareyou'

Figure 4: Core opening sequence in a telephone conversation.

The participants have individual knowledge of the core opening sequence in a telephone conversation. The expected structure of the core opening sequence mediates the interaction at an anticipated point of interaction – the opening of a telephone conversation -- making it easier for them to more effectively develop an intersubjective space within which to proceed with their telephone conversation.

Both actors have knowledge of the core opening sequence. Each utterance serves a dual function: it communicates content and it helps the participants to align their private representations of the core opening sequence. When a secretary in the office picks up the phone and answers "This is the computer science department", the content of his utterance identifies the receiver of the call, and it also marks the progress of the participants through the opening core sequence.

Suppose Simone wants to call her lover Jean Paul on the phone to let him know her flight has been canceled and tell him the flight number and arrival time of her new flight. In this situation, Simone has at least two internal frames that are mediating her behavior.

1. F_i – Simone's internal representation of the core opening sequence
2. F_j – Simone's internal representation of her plan to inform Jean Paul that her flight is canceled.

Simone dials Jean Paul's number. Each utterance in the conversation conveys content, but it also marks the progress of the interactants as they proceed through the core opening sequence. A deviation from either the frame F_i or the frame F_j will result in a breakdown.

Coordinating Representations

The term *representational system* refers to the entire collection of representational devices and content available at the scene of an activity (Hutchins, 1995ab). A representation system has three parts:

1. A set of representational media available to the participants.
2. A set of internal or external, private or shared, representations, including those provided in the design of the task environment and ones created at runtime.
3. A set of procedures for communicating, recording, modifying, transcribing, and aligning multiple, partial representations of the shared context.

In the classroom, during a lecture, the slides from the teacher’s presentation, what was written on the chalkboard, and the student’s notes are all part of the representational system in which the participants cooperate.

A design problem is to arrange the representational system to best support the expected recurrent interactions among the actors at a specific locale. Some subset of the design of a representational system is specifically engineered to support the co-construction of an intersubjective space among a group of collaborators.

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 at a recurrent point of coordination.

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

The stop sign is a coordinating representation that mediates the interaction among the participants at an intersection in the road. A clock in the classroom is a coordinating representation that helps to align private representations about when the class begins and ends. An appointment slip helps a patient to return to the dentist's office on the right day at the right time. A mail order catalogue helps the customer and the sales office reach agreement on purchase items, sizes, and prices. Tax forms help to coordinate citizens and IRS personnel in their efforts to exchange information. At the airport a passenger’s printed itinerary, the departure monitor, signs identifying the JetBlue™ ticket counter, and baggage claim tickets are all examples of coordinating representations that have been designed into the environment.

All artifacts can be used to mediate the co-construction of a shared understanding, but not all artifacts are designed to do that. A chair could mediate a point of coordination, but the chair was not designed with that purpose in mind.

Artifacts have both a tool and sign function (Vygotsky, 1978: p54-55). The tool function makes it easier to accomplish some task. The sign function effects how we think about the task. For a coordinating representation the sign and tool function coincide: the tool function of a coordinating representation is that it is a sign designed to mediate an interaction.

Not all external representations are intended to mediate an interaction, and therefore not all external representations are coordinating representations. A photograph is not a coordinating representation. The earlier drafts of this paper helped me to work out what I want to say, but they were not coordinating representations. A scratch piece of paper that is used to do multiplication problems is not a coordinating representation. A personal diary is not a coordinating representation, even if somebody other than its author reads it.

At many locales, media is available, like the whiteboard, which the participants can use to construct external representations that coordinate the activity of a group. These representations become coordinating representations, in the sense that is meant here, only if their usage continues beyond a single episode of cooperation.

In the ideal case, a coordinating representation effectively mediates the point of interaction it was designed to support. Non-normative senses of a coordinating representation can also emerge at a point of coordination. Ludwig sees the stop sign. He believes it is relevant to a point of coordination that he is anticipating. The frame F_j that is mediating his behavior grounds the sense he makes of the coordinating representation, and the sense he makes of it could be: he sees it, he knows it is not his turn, but he is going to drive through the intersection next anyhow. If Wolfgang sees the stop sign, the sense he makes of Ludwig's behavior may fit under the banner of "dealing with impatient drivers at stop signs". In any event, whether it be the normative or non-normative case, the coordinating representation is a powerful constraint on the intersubjective space the participants co-construct at runtime.

A Case Study

The study presented in this part of the paper documents the parts of the model that have to do with points of interaction and mediation. Subjects are organized into teams that are required to collaboratively achieve a set of goals. Each team works with one of two representational systems to co-construct an intersubjective space in which to operate as they achieve there collaborative aims. The representational systems have a chronological order: one of the representational systems includes coordinating representations that were specifically designed to mediate some of the points of coordination that recurred while using the first representational system. This set-up enables us to consider, in detail, a sequence of related interactions, within and across episodes of cooperation, such that continuity and change can be observed.

The data shows that the anticipation of certain problematic points of coordination result in the construction of conversational structures to mediate those points of coordination thereby reducing error and work. At a recurrent point of interaction, a conversational structure that emerged for the first set of groups does not develop if a coordinating representation is available that is a more effective form of mediation: conversational structure and coordinating representations have similar function. Some coordinating representations reduce the number of recurrent points of coordination. If adopted the use of a coordinating representation reduces the representational work for the actors and the amount of conversational interaction, while increasing the productivity of the group.

The data is collected from an online study of subject's collaboratively problem-solving in a simulated environment called VesselWorld. There are two versions of VesselWorld that will be discussed. The base version of VesselWorld will be used to show how the subjects invent conversational structure in response to breakdowns in anticipation of future points of coordination. The second version of the system, VesselWorld+, includes coordinating representations that were designed by the experimenters to mediate certain recurrent points of interaction.

Methods

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 of these technologies to collect data of the sort that was needed to study the dynamic of cognition and intersubjectivity. 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 have very high transcription costs, and they 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, within a community, of the intersubjective space that is constructed at runtime.

A key component of the VesselWorld environment was that all the interactions among the subjects were captured automatically in a replayable form. Every mouse click, every event, every shared bit of information, was recorded without bias within the transcript for a session. Moreover, the transcript automatically included markings for different types of events -- for example, a "planning event" or a "chat event". Because events are automatically marked on the transcript as being of a certain type, it is significantly less time consuming to analyze the data. The analyst can search through the data using any number of criteria: he can move forward to the next communication, round, plan action, or other such action within the system. See Landsman et al (2000) for a further discussion of this technology.

We have built several versions of VesselWorld and collected over a hundred hours of data. Several formal studies comparing teams of subjects using different versions of the VesselWorld platform have been done. This paper will compare only two versions of the representational system of VesselWorld: the basic system (VesselWorld) and VesselWorld+.

VesselWorld

In VesselWorld, three users, situated at three physically separate locations, engage in a set of cooperative tasks that require the coordination of behavior in a simulated environment (see Figure 5). In the simulated world, each participant is the captain of a ship. Their joint task is to find and remove barrels of toxic waste from a harbor and load them onto a large barge. Two of the users operate cranes that can be used to lift toxic waste from the floor of the harbor. The third user is the captain of a tugboat that can be used to drag small barges from one place to another. Segments of activity are divided into *rounds;* it takes at least 6 rounds of activity to move from one end of the harbor to the other. During a round of activity, participants plan out their future actions explicitly, and then submit them to the system. Once a participant has submitted her next action, she can no longer change it. When all three participants have submitted actions, the round ends, the system updates the state of the world, and the next round begins.

There are many complications in clearing the harbor. Some barrels are large and require the two cranes to join together and lift them simultaneously. What equipment is needed to retrieve a particular barrel can only be determined by the Tug operator, and only when he is next to it. The participants have limited (and non-identical) areas of perception, and the harbor must be searched to discover the toxic waste.

- Goal:
    - Remove all barrels of toxic waste from harbor and load them on the large barge.
- Domain objects:
    - The harbor
    - Barrels of toxic waste (large, medium, or small)
        - Can require special equipment to remove (dredge or net)
    - Small barge
        - Can be used to carry several barrels of toxic waste.
        - Can be moved from one location to another
    - Large Barge
        - Fixed location in the harbor.
    - Special equipment
        - Dredge
        - Net

- Participants:
  - Tug
    - Moves small barge
    - Examine barrel of toxic waste; determine if special equipment is needed
  - Crane1
    - Can lift small and medium size wastes.
    - With Crane2 can jointly lift larges wastes.
    - Special equipment: Dredge
  - Crane2
    - Can lift small and medium size wastes.
    - With Crane1 can jointly lift larges wastes.
    - Special equipment: Net

Figure 5: Overview of VesselWorld participants and task.

Most of the participant dialogue is centered on the barrels of waste and how effort can be coordinated in removing the barrels from the harbor and transporting them to the large barge. The participants must also keep track of what areas of the harbor have (or have not) been searched. The participants must discover and then keep track of the location of wastes. Initially this is the location of a waste in the harbor; later this includes whether a waste has been moved -- and if so, where -- and if it is on a small barge, in what order it was stacked. References to the wastes must be shared; these references can change depending on the circumstances.

A portion of the interface for the *base system* of VesselWorld is shown in Figure 6. The WorldView (the large window in Figure 6) graphically represents several kinds of information about the location and status of objects, from the perspective of an individual participant. It depicts the harbor from the participant's point of view; only a limited region of the whole harbor is visible at any one time – the shaded region in the figure. When two or more vessels have overlapping radiuses of perception, the participants can "see" each other to the extent that they know the other vessel(s) are nearby, but there is not sufficient detail to determine in what sort of activity the other ship is engaged. The participants can "mark" their map with labeled markers, but they cannot see each other's markers.

A second window of information is used for editing and displaying the user's current plan. A third window allows the user to access more detailed information about visible objects. Textual chat is used as the primary method for participants to communicate with one another during their cooperative activity; Figure 7 gives some examples of the kind of coded chatting in which the participants engage.

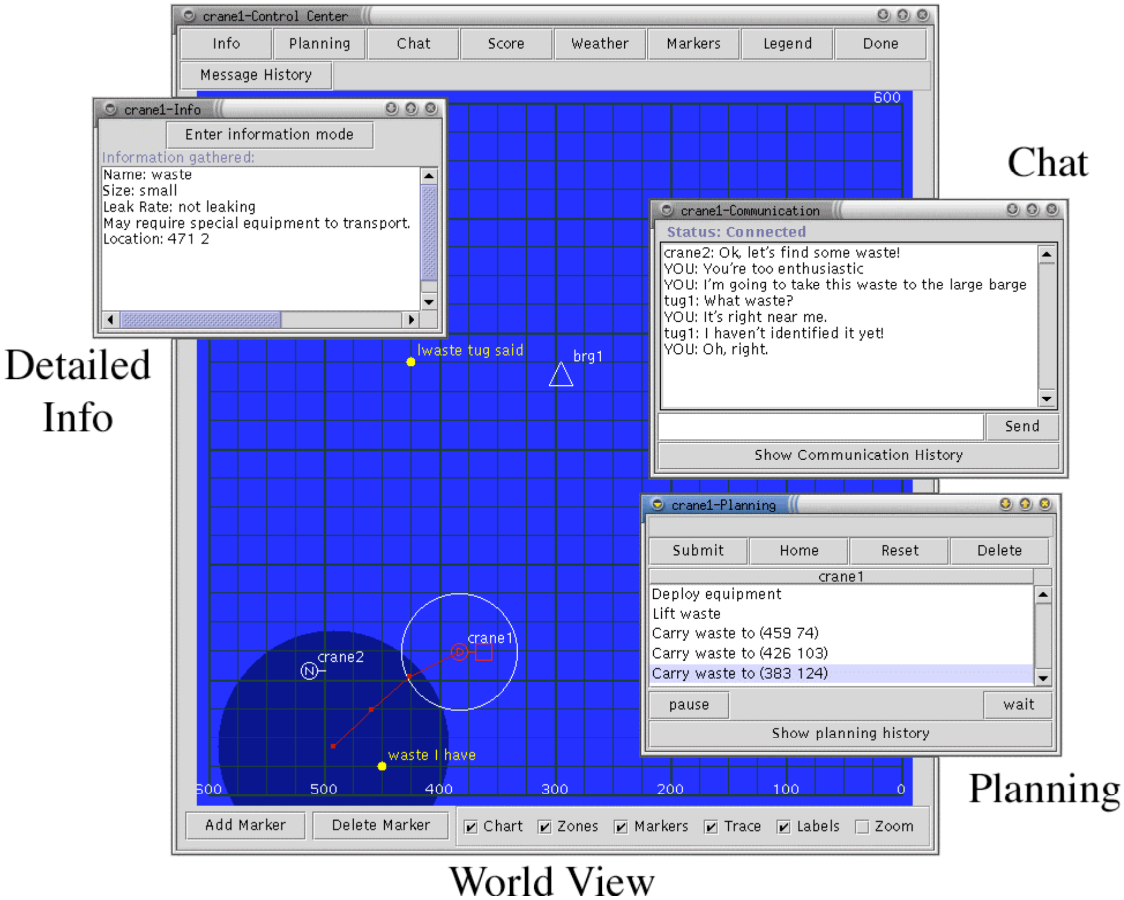

Figure 6: The interface for the basic system.

- Waste at 554,41 is small, dredge
  554,41 -> location of waste in harbor ; size of waste is small; special equipment (a dredge)
- New XL! At 200 431
  XL -> waste is extra large ; location is 200 431
- I guess I'll sweep the bottom, west to east
  Looking for wastes along the bottom of the harbor moving west to east
- [all] remove marker at 100,425
  all -> address to other two participants ; waste at location 100,425 has been removed
- 8 clicks for me to hit BB ; *It will take 8 rounds of activity to reach the* ***b****ig* ***b****arge.*
- I'll grab sX at 500 275 ; *participant will remove small waste (sX) at location 500 275*
- I have a leaker on my hands ; *barrel of waste is leaking toxic material into the harbor*
- I'm waiting at SW corner ; *participant is waiting in the SW corner of the harbor*
- w8ting
- killed sX at 500,275
- I still and 2 or 3 moves till I get there

Figure 7: Examples of chatting

VesselWorld+

The second version of VesselWorld (VesselWorld+) included three coordinating representations. The subjects using VesselWorld+ are sometimes referred to as the CR groups to indicate they had access to the additional coordinating representations.

The coordinating representations we developed for VesselWorld+ resulted from an analysis of a pilot study that was conducted with the base system. Our analysis of the pilot study data focused on the recurrent activities of subjects, especially those where the construction of a shared context was problematic or error prone. One important indication of trouble spots were those places where the participants invented conversational structure to organize their interaction. A discussion of the methods used for analyzing the online interactions can be found in Feinman & Alterman (2003).

Of the three coordinating representations we added to the VesselWorld environment, two were adopted by our subjects:

1. A coordinating representation that mediates the participants' efforts at points of coordination during a sequence of tightly coupled actions.
2. A coordinating representation that mediates the participants' efforts at points of coordination over shared domain objects.

A third coordinating representation was designed to allow the users to manage multiple plans. The idea was to create a coordinating representation mediated the subjects interaction as they rapidly sketch a high-level plan to manage multiple open tasks. The high-level planning coordinating representation was not used by any of the CR groups. The participants did not use the high-level planning window because the extra representational work needed to construct a representation of a high-level plan was not warranted. An analysis of the discourse showed that the plans the subjects created had a relatively short period of average relevance. Thus information about the plan was readily assessable from the short history of prior chat that was already available in the chatting window.

A formal study comparing VesselWorld to VesselWorld+ found several significant differences within and between the representational systems of VesselWorld and VesselWorld+; see Feinman & Alterman (2004) for further details. In this study each team of subjects was trained for two hours and then solved VesselWorld problems for approximately ten hours. We compared differences in performance from the first five hours of problem solving sessions to the second five hours for each team; after five hours the performance of the teams had stabilized.

All teams, regardless of which representational system they used, saw significant decreases, over time, in the number of chat lines they produced and the elapsed clock time it took to achieve their goals: the participants talked less and took less time to accomplish their tasks.

Between groups with different representational systems, there was a small reduction in the rounds of activity, but it was not significant. The use of coordinating representations, however, did result in significant improvement in performance by other measures of performance:

1. A 61% reduction in domain errors (errors in performing domain actions which led to a toxic spill)
2. A 57% reduction in the number of chat lines
3. A 49% reduction in clock time
4. A 38% reduction in system events (mouse clicks, etc.).

A reduction in domain errors indicates improvement in the runtime construction of intersubjective space. Because there was no significant improvement between groups in the number of rounds it took to complete a task, each of the other three findings – reductions in the number of chat lines, clock time, and system events – are also likely to be indicating that the intersubjective space in which the VesselWorld+ subjects operated was richer and easier to produce. Shown below is an analysis of the transcripts of VesselWorld and VesselWorld+ sessions that examines in detail the interactive work of the participants at some recurrent points of coordination that existed for both groups of subjects. The analysis documents how and why the coordinating representations produced positive change in the runtime production of intersubjective space.

Close Coordination

Tightly coupled actions involving the manipulation of large and extra large wastes require close coordination. If one crane starts lifting before the other crane, the waste will spill and leak toxic materials into the environment. In order to successfully lift, carry, and load on the barge an extra large waste, the participants must successfully negotiate several points of coordination:

1. The cranes must intend to cooperatively lift the same waste.
2. Both crane's must be in close proximity of the waste.
3. One of the cranes may need to deploy equipment necessary before the lifting begins.
4. During the same round of activity, the cranes must join together.
5. During the same round of activity, the cranes must jointly lift the waste.
6. During the same round(s) of activity, the cranes must jointly carry the large waste to the barge. (If the small barge or large barge is nearby this step is not needed.)
7. During the same round of activity, the cranes must jointly load the large waste onto the small barge.

These actions must occur in a strict order and must be synchronized. Errors in coordination result in result in a spill of toxic waste into the harbor.

The Base Group Invented Conversational Structure to Mediate the Interaction

As they prepared to do a joint lift, the subjects using the base system could "see" each other, but their perceptual capabilities were not sufficient to see what the other actors were doing in any detail. Each participant had a plan, but the participants could not see each other's plans. The major tool they had for co-constructing intersubjective space was the chat. The problems inherent in jointly lifting or moving a large or extra large toxic waste made for a recurring source of difficulty.

The base groups invented a conversational structure to organize operations on large and extra large wastes at each point of coordination. A set of *adjacency pairs* (Schegloff & Sacks, 1973) were used by the subjects to mediate the private understandings of these tightly coupled actions.

The first part of the adjacency pair was for one actor to propose to take a given joint action on the next round. The second part of the adjacency pair was for the other actor to confirm that he would take the corresponding action. So, if Crane1 proposes to do a joint load, Crane2 can confirm. For joint actions requiring multiple steps, each of the steps is proposed and confirmed using the adjacency pair structure.

Figure 8 shows a sample of dialogue where the participants used adjacency pairs to coordinate the handling of a large barrel of toxic waste. At 1 and 2, after jointly lifting a large barrel, Crane1 and Crane2 agree to do a joint carry followed by a joint load onto a barge. It will take three moves to reach their destination. In lines 3, 4, and 5, they tell each other they submitted their first move. At 8 the tug suggests a convention to simplify coordination. At 9 and 10, Crane1 and Crane2 tell each other they are ready to do the second part of the move. At 14, Crane1 states she is doing the third move. At 15-18 they plan and then they submit actions to do the joint load. At 19 and 20, they celebrate. Because the conversation of the users is mediated through textual chat, adjacency pairs do not strictly speaking occur one after the other; their positioning sometimes depends on the typing speed of the users. Other kinds of comments may end up interposed along the way.

1. Crane1: now a joint carry, clicked at 375,140 got 3 carrys
2. Crane2: i will do same
3. Crane2: move to first location
4. Crane1: submitted first
5. Crane2: ditto
6. Crane1: again?
7. Crane2: yes
8. Tug1: do you want to just type something in after submitting each turn
9. Crane1: submitted second
10. Crane2: ditto
11. Tug1: just some shorthand or something, for everyone so we know whats going on
12. Crane1: submitted third
13. Tug1: submitted
14. Crane2: submitted third
15. Crane2: Crane1: load, and then i'll to the same
16. Crane1: submitted load
17. Crane2: ditto
18. Tug1: submitted move
19. Crane2: hey, i think that worked!
20. Crane1: looks like it's Miller time. I think we did it.

Figure 8: A conversational structure.

After this conversational structure became a part of the group's common knowledge, only some of the progress of the interaction it produced was specifically marked in the chat window.

VesselWorld+

One of the coordinating representations available to the CR groups was designed to create a shared representation that would simplify the process of aligning private representations of time (see Figure 9). The shared planning CR enables a user to compare his projected actions to those of the other participants. For each participant, the next few planned steps are displayed in a labeled column for each participant. The actions are listed in order from top to bottom. (So, the next few planned steps of Crane1 are to deploy equipment and then lift some waste.) This representation provides an alternative method to the adjacency pair structure adopted by the base groups for mediating the interaction at points of coordination during tightly coordinated activities like lifting an extra large waste.

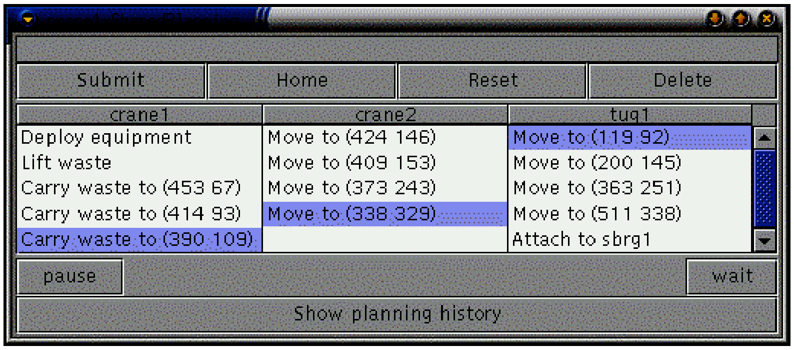

Figure 9: Timing of joint actions.

The CR groups used this shared external representation to mediate closely coupled actions. In order to submit an action to the system the users needed to add it to their plan anyway. So, from the point of the view of the users who had access to the shared planning window, having to talk about their immediate plan was just extra representational work. On more than one occasion it was observed that one crane would adjust his plan to match the plan of the other crane within the same turn, without any discussion. Removing the articulation costs of chatting reduced the cost of constructing intersubjective space.

Figure 10 shows the cranes' entire conversation while jointly lifting a barrel of waste and loading it on a small barge (compare this to the conversation in Figure 8). During round 7, the cranes and the tug all agree to remove wastes G and B from the harbor. During round 8, the Tug makes a joke, which Crane1 responds to during round 10. During round 11, the Tug announces that G1 is loaded.

1. CRANE1: G, then B
2. CRANE2: Okay.
3. Tug1: sounds like a plan
4. TUG1:
5. -------- End of round 7 ------
6. TUG1: a man, a plan, a canal, panama.
7. -------- End of round 8 ------
8. CRANE1:: that was *last* time : -)
9. -------- End of round 10 ------
10. TUG1: G1, loaded.
11. -------- End of round 11 ------

Figure 10: Conversation while lifting and loading an extra large waste.

Where the base groups used a conversational structure to mediate the cooperation during a joint lift, the VesselWorld+ groups (see Figure 11) used a coordinating representation. By the end of round 7, all three participants have planned out their next few actions and these are visible to all of them via the shared planning window. Crane1 will wait while Crane2 deploys equipment. In the succeeding rounds of activity, the cranes will join to one another, will jointly lift the waste, and then will load it on the barge. Through all of these actions, the Tug will hold steady waiting for the small barge to be loaded.

Figure 11: The planning window at the end of round 7.

Shared Domain Objects

In VesselWorld there are several domain objects shared among the participants:

1. The dredge.
2. The net.
3. The large barge.
4. The small barge.
5. Each of the actors.
6. The harbor.
7. The barrels of toxic waste.

The rules of VesselWorld define which crane has the dredge and which crane has the net. The WorldView shows the location of the large barge. Early on in a problem solving session the tug collects the small barge, thereafter it is with him. Unless the actors have overlapping radiuses of perception, they need to tell each other where they are in the harbor; sometimes this is reported as how many rounds of activity it will take to get to some other place.

Both the harbor and the barrels of toxic waste required considerable interactive work on the part of the subjects to construct and develop intersubjective space.

For the harbor, the participants regularly update their assessment of how much of the harbor has been searched and cleared. When these points of coordination occurred, each of the participants took a turn, reporting ether what had (or needed) to be searched/cleared or confirming the assessment of other actors. For all of the base groups these kinds of conversational interactions regularly occurred. With one exception, VesselWorld+ groups also engaged in this kind of conversational interaction; the exception co-opted the object list CR (described below) to track the status of the harbor.

By far the most difficult task for the subjects using the base system was aligning their private views of the location, size, special equipment needs, and references for barrels of toxic waste. Each of the participants used markers to keep track of the wastes and their various properties, but the individual marker lists would inevitably diverge, resulting in breakdown and a conversational interaction to re-align private views.

An example of where a conversational interaction is needed to repair a misalignment is when one crane intends to pick up a waste only to discover it is not there. He immediately comments upon this to the other participants, whereupon one of the other participants responds that the waste has already been removed. Another case was when one subject would "discover" a waste that have been reported previously by another subject. To repair this problem the actors conversationally interact to determine if it was the same waste. A third example of repair work was when two subjects "simultaneously" found the same waste. Because the participants have limit perceptual capabilities and cannot point it takes communicative work to establish they are talking about the same barrel of waste.

In anticipation of the breakdowns that resulted from keeping separate representations of the wastes and their locations, one of the groups invented a conversational structure, a "marker check", which they used to periodically compare private representations (see Figure 12). At line 1, Crane1 proposes to do a marker check. There is some intervening dialogue and then at line 3, Crane1 produces a legend. At line 4 he establishes an ordering, and at line 5 proceeds to list all the wastes he thinks they have found "south of the equator". At lines 6-10 the Tug and Crane2 initiate some repair work. At line 11, Crane1 continues listing all the toxic wastes, their locations, and the equipment needed. At lines 15-22, the repair work initiated earlier in the dialogue continues until all the differences in the participants' individual assessments of the known wastes are resolved.

1. Crane1: [ALL] ok I will dump all the markers ok with every1?
2. …
3. Crane1: Legend: [Sm|L|XL] – [Ni [no id'd] Net | Dr]
4. Crane1: from south east clockwise
5. Crane1: [Sm-Ni 50,0][Sm-Net 150,25][Sm-Ni 350,150][Sm-Ni 550,50] [Sm-Ni 600,100] That's all south of equator. NORTH coming up.
6. Tug1: 97441 and 72,368 already ID'd
7. Crane2: 350,150 is barge, isn't it?
8. Crane2: that's the problem
9. Crane2: stop dump I was there to… that's the Sbarge at 350,150
10. Crane2: confirm with TUG
11. Crane1: [xL-Ni 475,425][Sm-Ni 450,450][Sm-Ni 525,500][Sm-Ni 250,500][XL-Ni 200,475] [Sm-Net 100,425][Xi-net 75,375][Sm-Ni 25,575]
12. Crane1:
13. Crane1:
14. Crane1: –- END – Tug, confirm you have all those
15. Tug1: large barge at 400,325
16. Crane2: repeat: I say the sbarge ar 350,150, not a small waste
17. Crane2: [TUG] where are the small barges?
18. Crane1: [c2] – you sure? If you saw that, ok ill cancel as waste
19. Crane1: of yeah tug can tell us
20. Tug1: 350,150 is small barge
21. Crane2: ok, 12 wastes and no problems, let's get back to work
22. Crane1: okdoke. Still sweeping west

Figure 12: A marker check.

VesselWorld+

The object list is a coordinating representation that mediates the efforts of participants to construct intersubjective space for shared domain objects, and it potentially mediates the interaction at any number of points of coordination. A list of objects (with relevant properties) allows users to more systematically keep track of objects in the domain**.** This information is visible to all users and can be edited by any user. When a user discovers a waste, he can note it in the object list using a point-and-click operation. Entries in the object list can be displayed on the WorldView as markers. All of the teams that had access to the object list used it to mediate their interactions.

Figure 13 shows a portion opening dialogue in a VesselWorld problem solving session where users had access to coordinating representations; in total there are 17 lines of chat before the first round of activity ends. This dialogue ensues before all of the participants have submitted their actions to the system for the first round of action. At line 1, Crane1 ecstatically declares that he can see an extra large waste. At line 2, the Tug expresses his "envy". At line 3, Crane2 expresses his excitement that he can see both an extra large and a large waste. The rest of the opening dialogue, an additional 13 lines of chat, is concerned with planning.

This example demonstrates how a CR can reduce the numbers of points of interaction. For the base group, because the participants keep separate representations, each time the discovery, or any other kind of information exchange about a waste occurs, there is a point of interaction. For example, if the tug report he found a barrel at a specific location, the other actors need to add a marker to their map for them to keep track of the fact it exists. With the introduction of the object list that point of coordination no longer exists.

1. Crane1: I got an XL!
2. Tug1: I got nothing, you luck basrstartd.
3. Crane2: I got an Xl and an L, mommy! ;)
4. Tug1: Merry christmas kids….
    -
    -
    -

Figure 13: Opening dialogue.

Figure 14 shows the object list that is constructed by the time all the participants have submitted their first action. Only three of the entries into the object list were explicitly mentioned in the opening dialogue, and none of these were explicitly named. Much of the dialogue that accompanied the discovery of a new waste in the groups using the basic system is now mediated by a coordinating representation.

⊕	Name	Location	Size	Equip	Action	Leak	Notes
☒	G1	556 465	XLarge	Unknown	Located	Not Leaking	
☒	A1	186 107	XLarge	Unknown	Located	Not Leaking	
☒	m	550 447	Small	None	Located	Not Leaking	
☒	A2	249 21	Large	Unknown	Located	Not Leaking	
☒	m	250 149	Small	None	Located	Not Leaking	
☒	m	449 349	Small	None	Located	Not Leaking	
☒	SB	305 310	XLarge	None	Located	Not Leaking	

Figure 14: The object list.

Discussion

The VesselWorld task is an artificial one, but it is by no means simple. As is the case of the real world, the behavior of VesselWorld participants is embedded in a rich representational system.

VesselWorld(+) is not a face-to-face interaction but such is the case for many real world interactions. There is, however, lots of evidence that much material relevant to the construction of intersubjective space is available only when the actors are physically present with one another. (The struggle to replicate face-to-face behaviors online is significant evidence that portions of the production of intersubjective space would be lost in a VesselWorld experiment.) On the other hand, breakdowns in collaboration and resulting efforts on the part of the subjects to produce verbal descriptions to compensate for what is missed does tell us something, and it is in a form that may be easier for the analyst "to see".

There is the makings of a community when actors have cooperative behaviors they recurrently perform within a representational system. Two shifts of clerks at the local supermarket to a certain extent are part of the same community by virtue of the fact that their work occurs within the same representational system. But there are lots of other important features of the supermarket example – for example identity and participation within a community -- that makes for community, is relevant to construction of intersubjective space, and that at best marginally exists for VesselWorld subjects as a community. Because the experimenters designed the CR's they are to a certain extent part of the VesselWorld community. That they are not directly involved in the collaborative work of subjects at runtime, or that they train the subjects on how to use the representational system, is analogous to some cases of real world design and training.

VesselWorld is more natural and real than many experimental tasks that have been devised. Switching to videotape is not an antidote to all the methodological problems of doing an online study. Adding any sort of recording device at the scene of an activity (in plain sight) can create an "on stage affect" to the performance of the subjects. If the subjects are unaware of the recording it would be natural, but also illegal. Collecting videotape data within and across multiple episodes of cooperation in a rich, and changing, representational system is prohibitively expensive, and analyzing all that data is mind bogglingly complicated. Collecting and analyzing data from multiple related episodes of online cooperation, however, is both affordable and doable.

Concluding Remarks

To achieve coordination between actors it is only necessary that the participants' mental representations of the situation are functionally equivalent. The methods the participants use to coordinate their behavior, generate, either intentionally or unintentionally, partial understandings, lack of understandings, and misunderstandings. Procedural framing of each actor's contribution to the production of intersubjective space is sufficient to cognitively model how the functional equivalence of mental representation enable coordination of cooperative behavior and also explicate how, for example, intersubjective space is co-constructed when the participants have a lack of understanding or even misunderstandings.

Because of dynamics, there are points of interaction and coordination where the participants must work to re-align their private views of the situation. The organization of conversation is a flexible and powerful enough medium to handle all kinds of breakdowns and contingencies. The anticipation of recurrent points of coordination, however, can result in the creation of additional structure that will more effectively mediate the expected interaction if and when it occurs.

Conversational structure and coordinating representations are two alternate means of more effectively orienting and coalescing individual viewpoints at expected points of coordination. Within a community, the accumulation of structure to mediate recurrent points of coordination is a significant feature of how the runtime production of intersubjective space adjusts and improves despite the dynamics of everyday behavior.

Collecting data for this kind of research is not a trivial task. Using online cooperation as a source of data solved a number of technical problems. The most important of these was that it enabled us to capture more than a single snapshot of the community at work. It allowed us to document interactive work and emergence of conversational structure within the confines of a single representation system and then to compare that to the progress of another group of participants with a different representation system that included some new coordinating representations. The data showed that the anticipation of certain problematic points of coordination resulted in the invention of conversational structure and that the invention of conversational structure and coordinating representations have a similar function. The data also showed that coordinating representations add to the richness and effectiveness of the intersubjective space in which the participants operate either by restructuring the interaction so that fewer points of coordination are required overall, or by reducing the representational work and conversational interaction that is needed at a recurrent of coordination.

Acknowledgements

This work was supported by ONR Contracts N00014-96-1-0440 and N66001-00-1-8965.

References

Agre, P and Chapman, D. (1990) What are Plans For? In P. Maes (editor) *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back.* Cambridge: MIT Press, 17-34.

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. (2001) Coordinating Representations in Computer-Mediated Joint Activities. *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 Schaefer, E. (1989) Contributing to Discourse. Cognitive Science 13, 259-294 .

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

Clark, H. and Marshall, C. (1981) Definite Reference and Mutual Knowledge. In Joshi, Webber, and Sag editors, *Elements of Discourse Understanding*, Cambridge University Press.

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.

Miller, G., Galanter, E., and Pribram, K. (1960) Plans and the Structure of Behavior, Holt, Reinhar, and Winston, Inc.

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. (1986) Plans and Situated Actions. Cambridge: Cambridge University Press.

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.