Coordinating Representations in Computer-Mediated Joint Activities

Richard Alterman, Alex Feinman, Josh Introne, Seth Landsman
Department of Computer Science
Brandeis University
Waltham, MA 02454 USA

Abstract

This paper develops, in the context of the interdisciplinary literature on coordination, the concept of a *coordinating representation* as an everyday method for structuring the coordination of actors engaged in a non face-to-face joint activity. Evidence is provided by applying the idea of coordinating representation to the development of a computer-mediated cooperative activity.

Introduction

A critical reasoning problem confronted by actors as they engage in their everyday activities is the maintenance of coordination (Clark, 1996). Within a community of actors, designs that organize (structure) behavior in recurrent situations of cooperation develop over time. Once developed, the expectation that a given sort of structure might be in place for a given kind of situation simplifies the interaction among the participants while reducing mental effort, physical work, and errors (Alterman & Garland, 1998). In non face-to-face interactions, structures that simplify the coordination of a conventional behavior are coded into a *coordinating representation*. The coordinating representation helps the participants to jointly make sense of the situation in the absence of a face-to-face interaction.

An everyday example of a coordinating representation is the "stop sign". The stop sign is a representation shared among the participants at a traffic setting. The stop sign presents a structure for organizing the collective behavior of drivers, pedestrians, and cyclists at a busy intersection. The interpretation of the structure imposed by the stop sign is negotiated during the activity. Things may run smoothly at the intersection - but there will also be interruptions. An impatient driver piggybacks on the driver in front of him. A pedestrian decides to ignore the stop sign altogether.

The first part of this paper will develop the notion of a coordinating representation in the context of the interdisciplinary literature on coordination. The second part focuses on the cognitive engineering task of building coordinating representations for computer-mediated joint activities. The last part of the paper presents an experimental evaluation of the utility and function of the coordinating representation.

The Problem of Coordination

Whether it is greeting someone, or planning a potluck dinner party, or moving through a doorway, or forming a queue at the coffee shop - there are always problems of coordination. When you greet someone, depending on the circumstance, you may say “hi", shake hands, slap hands, hug, kiss, or ignore. Each form of greeting (except the last) requires coordination (and cooperation) among the participants. For a potluck dinner party, the meal must be coordinated for taste, balance, and variety. The meal can include appetizers, main courses, desserts, and beverages; a preponderance of one or the other detracts from the meal. For many doorways, there is not enough room for two people (say, in conversation) to pass through the doorway shoulder-to-shoulder. To effectively move through the doorway the participants must coordinate on an order as to which one passes through the doorway first, second, ... and who is to hold the doorway open. The queue at the coffee shop begins and ends at a certain place; people line up in the order they arrive.

Some examples of coordination problems are the assignment of roles, the establishment of location, manner, and structure, and issues of sequencing; timing and co-reference.

Suppose Tipper and Al are re-arranging furniture in the house. Each of the above kinds of coordination problem may come into play as they move the old couch from the living room, down the stairs, around the corner, through a doorway into the basement. Al's role may be to back down the stairs holding the front of the couch; Tipper walks forward holding the backend of the couch. Initially they meet in the living room. Their path as they carry the couch begins in the living room and ends at the basement. Their manner may be slow and cautious, so as to avoid bumping into walls and doorways. At certain points they are tilting the couch at an angle so they can move down the stairwell without bumping the couch into the ceiling. Coordination at *the boundaries between phases of the activity* (Clark, 1996), must be jointly engineered by Tipper and Al as they shift from moving down the stairs to moving through the doorway. In order to move the couch down the stairs, Tipper and Al need to establish co-references for features of the stairwell (e.g., the low ceiling) or the situation (e.g., an unexpected problems they encounter). Some of the coordination problems are 'solved' before action begins (e.g., Al walks backwards and Tipper walks forwards); others are resolved as the action proceeds (e.g., the coordination problems entailed by the low ceiling in the stairwell).

The term *structure for behavior* is used here to refer to the kinds of information exchanged between Tipper and Al in order to achieve their joint task and maintain coordination - examples of which are the assignment of roles, the path, the manner… Not all the information exchanged is a structure for the current behavior. For example, Tipper and Al are also socializing as they proceed with their activity. Nor are all structures for joint behavior exchanged at runtime: both Tipper and Al are likely to have prior experience at moving a couch through a doorway. Using both the social exchanges of information about structure and the recollection of prior related experiences, the participants must jointly reason out and construct a behavior which achieves their shared goal of moving the old couch from the living room to the basement.

The structures relevant to a given act in the current activity that are available before the act may be either recalled, planned, the result of an explanation, or designed. Both Tipper and Al may remember previous occasions when they moved furniture. For the difficult portions of their task, they may explicitly create a shared plan (Grosz & Sidner, 1990), an agreed to structure - you do this and I'll do that - for the behavior. If the structure for behavior is produced after a given behavior is completed, it is called an explanation (Mitchell, et. al., 1986), which can become realized in future related episodes of joint activity. Over time, for joint activities that Tipper and Al regularly do, behaviors become conventionalized and designs for the structure of those behaviors will begin to emerge (Alterman & Garland, 1998).

As Tipper and Al perform their activity, the fact that they are co-present allows them to monitor the progress of their joint activity. Because they can see one another, they can use body position to communicate information. Throughout their activity they can speak to one another in order to co-develop, for example, a procedure for moving the couch down the stairway. Their comments to one another are exchanged without delay, in the course of their joint behavior. The actions that form their conversation and activity occur sequentially.[1]

Other kinds of joint activity do not allow for a face-to-face interaction, so other methods or mediums must be introduced to support the exchange of structural information. Performance depends on the participants communicating - by these altered means - information relevant to design, plan, and commitment. For computer-mediated tasks, the trick will be to convert structures (designs) that are naturally produced in conversation by the users into external representations that can mediate similar sorts of cooperative activities in the future. The design of the external representations that are developed will focus on simplifying the most difficult coordination problems that typically confront users.

The Coordinating Representation

A coordinating representation is an external representation shared among participants in a joint activity. It is designed for the activity-at-hand and reduces the complexity of the coordination task. It mediates and structures the activity. It has the designated purpose of helping participants to achieve coordination in non face-to-face cooperative activities. Its meaning is based on conventional interpretation. It signals to the participants - without dictating action - that a convention of behavior is in place.

Consider the scene at the airport. For the passenger, the printed itinerary that her travel agent sent her helps her to stay coordinated. The itinerary identifies her flight destination and number. When she arrives at the airport, she uses the listed flight number to select among the flights and gates listed on the departure monitor for American Airlines. The design of the destination monitor (first listed in alphabetical order of destinations and then by time of departure) reduces her cognitive load in finding the departure gate for her flight. When it comes to finding her departure gate, the itinerary and the departure monitor are two coordinating representations that help to replace a face-to-face interaction with a mediated one.

Alternately, suppose the passenger needs to "check in" some luggage before proceeding to her gate. What coordinating representations are used to insure her bag makes the trip? Now, upon arrival at the airport, the passenger looks for the check-in counter for the airline from which she purchased her ticket. Large signs displaying airline logos indicate where each airline is located. Smaller signs divide the queue into first class and regular passengers. As the passenger puts her bag on the scale, the clerk attaches a tag indicating airline, flight destination, and flight number. Later, a bagger must transport in a truck the bags to the cargo space of the plane. A *complex sheet* that links flights to destinations and unique aircraft identification numbers is used by the bagger to achieve his goal (Goodwin & Goodwin, 1996). The organization of the complex sheet makes the access of information more efficient.

Each of the coordinating representations used to get both the passenger and her luggage on the correct plane has both a social and an individual function. From the perspective of the social, the coordinating

[1] This list is adapted from an analysis developed by Clark & Brennan (1991) to explicate differences among various kinds of mediated communication.

representation preserves a set of references for objects shared among the participants. From the perspective of the individual, the coordinating representations simplify access to the information that is being exchanged.

There are many other examples of coordinating representations in everyday life. 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....

Experimental Platform: VesselWorld

For the last several years we have been building a same time/different place groupware system (VesselWorld) as an experimental platform for analyzing real time computer-mediated collaborations. A demo of the system was run at CSCW 2000 (Landsman et. al., 2000).

There are several important characteristics of the joint activity of participants in a VesselWorld problem-solving session. Participants have different roles (both predefined and emergent). Cooperation and collaboration are needed to succeed. Participants must develop a shared understanding of an unfolding situation to improve their performance. Uncertainty at the outset makes pre-planning inefficient in many circumstances. There are numerous problems of coordination.

In VesselWorld, there are three users engaged in a set of cooperative tasks that require the coordination of behavior in a simulated environment. In the simulated world, each participant is a captain of a ship, and their joint task is to find and remove barrels of toxic waste from a harbor. Two of the users operate cranes that can be used to lift toxic waste from the floor of the harbor. The third user is captain of a tugboat. The cranes are able to individually lift and carry small or medium toxic waste barrels, jointly lift large barrels, and jointly lift (but not carry) extra large barrels. The tugboat cannot lift barrels, but can attach to, and move, small barges. Small barges may hold multiple barrels. Each captain has a small radius of perception. Many barrels require the use of other equipment in addition to the cranes. The tugboat captain is the only one who can examine barrels to determine equipment needs. Barrels can be leaking - or will begin to leak if they are dropped - in which case the leak must be contained by the tug.

The VesselWorld interface provides to each user several different windows of information. The World View (not shown) depicts the harbor from the point of view of a participant, who can only see a limited region at one time. The World View graphically represents several kinds of information about the location and status of objects from the perspective of an individual participant. A second window of information is used for planning. A third window allows a user to access more detailed information about visible objects. A chat window allows participants to communicate with one another using an electronic chat.

In a base version of the VesselWorld system, participants can only coordinate by electronic chatting. Most of the participant dialogue is centered on the barrels, and how effort can be coordinated in removing the barrels from the harbor and transporting them to a large barge. During a problem solving session, the flow of information between participants using the base system is continuous. It is the responsibility of each actor to add information conveyed to him by another actor to his or her private representation (either by taking notes, marking the map, or remembering), or be prepared to examine the history of chatting at some appropriate future time. Any information that is lost, misunderstood, never recorded, or never transmitted in the first place, can lead to discrepancies between the participants' individual assessments of the situation.

An analysis of participant dialogue determines a set of problem areas in organizing behavior in relation to a shared domain object. So, for example, a large volume of information must be exchanged over the naming, status, location, and properties of the toxic wastes. In a second version of the system, coordinating representations are introduced that basically structure and simplify the exchange of information in the problem areas of coordination.

Analysis of Electronic Chatting

The electronic chatting amongst participants is used as a basis for developing some coordinating representations. As the analyst reviews the discourse, she needs to look closely at using coordinating representations to simplify the most common interactions, fix repeated errors in coordination, and replace conventions developed by users during the course of a problem-solving session. The goal is not to entirely replace other forms of communication with coordinating representations. Rather the analyst wants to use coordinating representations to improve performance - thereby simplifying the interaction - at critical points in the ongoing cooperation among participants.

The analysis was framed by cognitive work on the problem of coordination that was presented at the beginning of the paper. Figure 1 shows a list of the kinds of methods that were used by participants to coordinate their joint activities. The participants did some planning by assigning roles or agreeing to sets of actions. During the activity, a fair amount of chatting was used to initiate joint actions that were tightly coupled; for example, to lift an extra large waste, the cranes have to begin lifting during the same time segment. Also found in the discourse were examples of the participants creating conventions to simply the exchange of information for recurrent problems of coordination. Chatting was continuously used throughout each session to establish references and exchange information about shared domain objects.

- Plan (provide orientation: delimit tasks)
  - Plan to do; Role assignments
- During Activity (Entry & Exit into Phases)
  - Synchronization; sequencing; step; turn-taking; Action taken; See; Initiating Statement
- Develop conventions
- Co-Referencing and the exchange of information
  - Refer to status, location, feature, identity of object

Figure 1: Taxonomy of coordination methods.

Figure 2 shows a sample dialogue of the kind of close coordination users needed to do in order to time closely coupled activities. 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 21 and 22, they celebrate.

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 2: A sample dialogue

Three Coordinating Representations

The analysis of the pilot study discourse identified three recurrent areas of coordination activity:

1. Timing of closely coupled activities
2. Establish references for, and exchange information about, shared domain objects and their status.
3. Higher-level planning to manage multiple cooperative activities

None of these should be surprising as possible areas of difficulty: each of these has been suggested by prior theoretical analysis. But there are also other potential problem areas. So the problem for the cognitive engineer is to determine which things are problematic for the task-at-hand.

Some sketches of three coordinating representations were developed and later refined through an interview with one of the test groups in the pilot study.[2] After the interview, the iterative design process continued by a cycle of (re)design, implementation, and evaluation. The periodic evaluation came in several forms, including expert reviews, in-group experimentation, and study groups paid for at Brandeis University. What resulted from this process were three coordinating representations that were designed both to simplify the interaction among participants (the social part) and structure it so as to reduce the cognitive load of each user (the individual part) in her use of the mediating representation.

The coordinating representation showed in Figure 3 allows a user to compare his projected actions to those of the other participants. The next few projected steps of each actor is displayed in a labeled column for each participants. The actions are listed in order from top to bottom. (So, the next projected step of Crane1 is to do deploy equipment and then he will lift some waste.) Each user has control of only one column, his/her own. This representation improves timing on exit and entry of phases for tightly coordinated phases of activity by allowing participants to compare each other's next few projected actions.

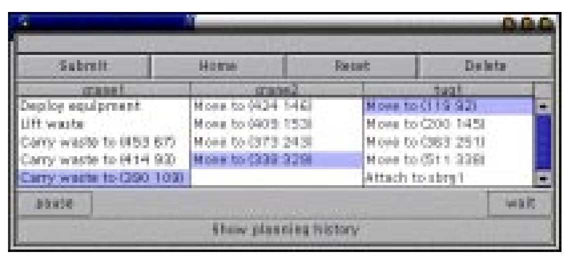

Figure 3: Timing of joint actions.

The second coordination representation is the *object list,* which contains a list of objects with relevant

---

[2] S. Kirschenbaum at NUWC collected the data for this pilot study.

properties in a table format. Columns provided information about the name, object type, location, and equipment needed for a given object. The organization of this information reduces the cognitive load for the individual, by organizing information relevant for decision making into predetermined representational structure.

A third coordinating representation was designed to allow the users to do *high-level planning*. The idea was to create a space where the participants could rapidly sketch a high-level plan that would help them to manage multiple open tasks. There are three columns in this window: one for each actor. Each column could be used, for example, to abstractly represent that each actor is currently searching a different part of the harbor. Further down each column, the participants could indicate that they are committed to a plan to move, in order, wastes 1, 2, and 3 onto a small barge. A palette at the top the window allows users to rapidly build a description of a joint action sequence. Actions are one of a small set of action primitives, i.e., MOVE, SEARCH, and CONTAIN. Color-coding of entries in the high-level plans allows participants to indicate both accomplished tasks and future commitments.

Experimental Evaluation

An experimental evaluation conducted at Brandeis compared the performance of teams of participants with (and without) the coordinating representation. Three groups could only electronically chat during problem-solving sessions, and three groups could chat but also had access to coordinating representations. Each team was trained and then played for about 10 hours over several sessions of problem solving. All events that occur during a problem-solving session are recorded in a log file by the system. A VCR-like device was used to review and analyze the decision making of each group. A more complete discussion and detailed analysis of the experimental data, with numerous examples, can be found in Alterman et al (2001).

Quantitative Analysis

One measure of general performance is the amount of clock time it took the participants to solve a problems:[3] there was a 49% improvement in clock time to complete task for those groups using coordinating representations. Another measure of user work indicates that there was a 38% reduction in the number of events generated while completing tasks of comparable difficulty. Because the coordinating representations pre-structure certain exchanges of information we expected to see a reduction in the quantity of electronic chats: there was a 57% reduction in the amount of electronic chatting. Because one of the coordinating representations dealt with commitment (high-level planning), another with timing, and a third with the exchange of information about equipment requirements for lifting barrels, we expected to see a reduction in domain errors: total errors were down 61%. However, a closer analysis of the data reveals that the high-level planning coordinating representation was used hardly at all. Further discussion of this last point is below.

Qualitative Analysis

For the groups that did not have access to coordinating representations, the predominant method for maintaining a common view of the world was for the participants to continuously *report* on their current activity via electronic chatting. One strategy for avoiding differences in assessment was to engage in a conversation to *review* the status of one or another of the shared domain objects. Whenever discrepancies in the assessment of a situation unexpectedly developed the participants engaged in *repair* work to re-mediated between alternate representations of "reality". Participants also regularly *confirmed* that somebody else's report or repair was received. Each of these techniques was important to the functioning of the groups using the basic system in maintaining a joint sense of their common enterprise. These groups also developed additional structures to simplify the exchange of information using the electronic chat window. The simplest of these were naming conventions. A second example was a set of conversational structures that were developed by each group to support coordination of closely coupled actions.

The general advantage of the coordinating representation was that it simplified the problem of establishing a consistent representation of the situation among the participants.

One advantage that accrued to the users who had access to the coordinating representation that supports the *timing of joint actions* is that it required no extra work on the part of the participants to build. 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 have access to the shared planning window, having to talk about their cooperative activity is just extra work. Another advantage was that one actor now has the opportunity to spot potential problems in another actor's plan.

Much of the dialogue that accompanied the discovery of a new waste in the groups using the basic system was mediated by the *object list* for the groups that had

[3] These results have 95% confidence intervals and are normalized for the complexity of the problem. Problem complexity is a weighted sum over all wastes taking into account size, equipment needed, and distance from large barge for each waste.

access to coordinating representations. Identifiers were attached to each of the "objects" that were found. Pointing and clicking was used to add entries to the object list, thus precise locations for each of wastes that were found could be stored. These aspects of the object list simplified the process by which the actors established references and referents. Because the object list was a shared representation, much of the consistency checking that the users of the base system had to engage in was no longer necessary. Rather than having two private representations that must periodically be reconciled by electronic chatting, the users could share a single representation. This scheme reduced the number of conflicts between different conceptions of the shared workspace, but it also eliminated the work involved in re-mediating discrepancies between alternate views of the shared domain objects.

The high-level planning window was not used by any of the groups. The surveys we collected from the subjects show that the chief problem with the high-level planning windows was that, given the rewards it provided, it required too much work to complete. Further analysis shows that the problems that the high-level planning window was designed to fix continued to occur.

We are developing two solution paths to fixing this problem. The first is to do a better job of modeling the work of the individual user in cooperation with the other users (Feinman & Alterman, 2001). The second approach is develop some AI techniques that would allow the system to fill out portions of the high-level planning window semi-automatically (Introne & Alterman, 2000).

Concluding Remarks

The overarching interest of this research is to continue to develop a framework for Cognitive Science that depends not only on the mental operations of the individual but also on the social interaction within which it is embedded. (An underlying thesis is the cognition is irreducibly social.) Application domains involving the computer-mediation of joint activity are significant areas of research because they allow one to investigate both the social and individual aspects of cognition. The methodology that was used for developing coordinating representations for VesselWorld reflects these commitments and attitudes. There are two parts to the methodology: a social and an individual one. During the social part, the developer can collect data on the usage of the base system and do an analysis of the information exchanged among participants (a discourse analysis) that helps them to stay coordinated. A key is to identify recurrent problems of coordination that showed up in the pilot version of the system. During the individual part, the designer tunes the initial approximations for coordinating representations to the cognitive operations of the individual user. During this phase, representations are iteratively designed to simplify the work of the individual user in creating and accessing the coordination information that is shared among the participants.

Acknowledgments

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

References

Alterman, R. and Garland, A. (1998). Convention in Joint Activity. TR# CS-98-199, CS Department, Brandeis University (1998). To appear in *Cognitive Science* 25(4), 2001.

Alterman, R., Feinman, A., Landsman, S. and Introne, J. Coordination of Talk: Coordination of Action. TR# CS-01-217.

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

Clark. H. and Brennan, S. (1991). Grounding in Communication. In Resnick, Levine, and Teasly editors, *Perspectives on Socially Shared Cognition* (pp. 127-0149)

Feinman, A., and Alterman, R. (2001) Modeling Communicative Behavior in a Groupware System. TR CS-00-211 Computer Science Department, Brandeis University. To appear in HCI 2001.

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

Grosz, B. and Kraus, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, 86:269-357.

Grosz, B. and Sidner, C. (1990). Plans for discourse. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, I*ntentions in Communication*, pages 417-444. Bradford Books, Cambridge, MA.

Introne, J. and Alterman, R. (2001) Segmenting Usage Data in Collaborative Systems. TR CS-01-215. Appeared in workshop on "Dealing with Community Data" at CSCW 2000.

Landsman, S., Alterman, R., Feinman, A, Introne, J., VesselWorld and ADAPTIVE, TR CS-01-213, Brandeis University. Demo given at CSCW 2000.

Lewis, D. (1984) *Convention: A philosophical study.* Harvard University Press.

Mitchell, T. Keller, R. and Kedar-Cabelli, S. (1986) Explanation-based generalization: A unifying view. *Machine Learning*, 1:47-80.

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