Research and Practice in Technology Enhanced Learning
Vol. 4, No. 3 (2009) 281–305

USING A CONTINGENCY GRAPH TO DISCOVER REPRESENTATIONAL PRACTICES IN AN ONLINE COLLABORATIVE ENVIRONMENT

RICHARD MEDINA* and DAN SUTHERS†

Department of Information and Computer Sciences
University of Hawai'i at Manoa, 1680 East West Road
POST 317, Honolulu, HI 96822, USA
**rmedina@hawaii.edu*
†suthers@hawaii.edu

People implicitly negotiate use of representations during learning, even in distributed online settings, but due to the temporally and spatially distributed nature of interaction, special analytic tools are required to uncover the development of representational practices in such settings. In this paper, we show how logs of online activity can be analyzed using specialized tools to recognize patterns in the participants' use of representations and show how negotiated representational practices affect how learners collaborate and influence each other.

Keywords: Representational practices; interaction analysis; online collaboration.

1. Introduction

Shared representations — constructed, manipulated and interpreted by participants — both influence and are appropriated by participants in the course of their collaborative interaction (Roschelle, 1996; Suthers & Hundhausen, 2003). Representational practices (e.g. how inscriptions are interpreted as representations, and role specialization with respect to construction and maintenance of these representations) are implicitly negotiated through cycles of innovation, adoption and revision (Danish & Eneydy, 2006; Dwyer & Suthers, 2006; Shipman & McCall, 1994; Stahl, 2006). Although much research on representational practices has been undertaken in face-to-face contexts, our data as well as others' (e.g. Overdijk & van Diggelen, 2008) shows that representational practices develop in online settings as well as face-to-face, and can even take place over extended periods of asynchronous interactions. The practices and roles so negotiated have implications for learning, as they can affect the extent to and ways in which learners collaborate and influence each other's views. However, it can be difficult to see implicit negotiations and their consequences when interaction is distributed over time, across workspaces, and over multiple modalities as is common in media-rich asynchronous online learning. To begin to address this problem, we have developed an abstract transcript format, the contingency graph, and tools for manipulating and visualizing this graph that are a step towards a toolkit for finding negotiated patterns of interaction and other relevant phenomena. A *contingency* is a way in which participants' actions are observably contingent upon prior actions and their artifacts, for example, through media dependencies, representational similarity, and semantic overlap, as discussed in Suthers, Dwyer, Medina & Vatrapu (2007, 2009) and later in this paper. A set of contingencies identified within a sequential data set form a contingency graph data structure.

This paper reports on how we used a tool for visualizing and manipulating a contingency graph in conjunction with screen videos to identify meaningful episodes of online activity that illustrate the negotiation of representational practices and how these practices led to specific observed outcomes in a collaborative problem solving session. We begin with a description of our analytic approach of identifying and interpreting contingencies. This is followed by a description of a tool we developed to support our use of contingency graphs. Finally, we illustrate this approach with a case study of dyadic interaction in an online asynchronous collaborative environment in which we utilized both a contingency graph generated from computer log files and video screen capture data taken from the participant machines.

2. Background

Computer generated log files of user interaction in shared online environments are commonly used as source data for the analysis of collaborative interaction (e.g. Bruckman, 2006; De Wever *et al.*, 2006; Larusson & Alterman, 2007; Martinez *et al.*, 2003). These machine readable histories of software events are amenable to computational methods for aggregating, searching, filtering, or visualizing sequential data in support of a range of analytical approaches (e.g. Aviv, 2003; Barcellini *et al.*, 2005; Hmelo-Silver, 2003, Landauer, Foltz & Laham 1998; Sanderson & Fisher, 1994). In our developing work on uptake analysis (Suthers, 2006; Suthers *et al.*, 2007, 2009), we have also begun to explore computational tools for analyzing collaboration and interaction from log files. Contingency graphs have been a useful starting point for these endeavors as we can generate some aspects of the graph directly from the log files. In this paper, we report on analyses in which we combined the use of such a graph with video analysis. We will present a prototype visualization tool constructed by the first author, the Uptake Graph Utility (UGU). The primary motivation for developing the UGU was to allow the analyst to control the visual rendering of a contingency graph representation in service of specific research questions or hypotheses. The next two sections provide a brief overview of contingency graphs followed by a description of the UGU prototype tool.

2.1. *Brief overview of contingency graphs*

The notion of a contingency graph has been useful in our work across a diverse set of data sources. The contingency graph supports our theoretical orientation and provides a consistent notation for generating analysis-specific artifacts such as the visualizations discussed in this paper. At one level of description, a contingency graph is a data structure for representing relations between events, particularly acts in which participants interact with media. At another level, a contingency graph reflects a theoretical commitment to the importance of the sequential organization of action and its contingency upon context, which implies that relationships between acts should be the basic units of analysis, not the properties of isolated acts. Our analytical methods are designed and deployed to recognize, document, and explain these relations in the diverse technology contexts in which they are situated. As illustrated in the case study presented later in this paper, the emphasis on the relation as unit of analysis provides a wider interpretive lens for understanding the range and complexity of distributed online interaction (Suthers, Dwyer, Medina, & Vatrapu, 2007). For present purposes, we will focus on the notational features of contingency graphs as they relate most closely to the visualization and analytic work presented here. We outline these graph features in the following two subsections.

2.1.1. *Vertices: events*

The events represented by vertices may include any manipulation of the medium that is available to participants, including (for example) not only the creation of media inscriptions (e.g. posting a message, making an object in a workspace), but also manipulation of those inscriptions (e.g. moving objects closer to each other) and perception of those inscriptions (e.g. opening a message to read it). The graph also records computer-initiated events such as the display of inscriptions that come from other participants in an asynchronous environment.

2.1.2. *Arcs: contingencies*

A contingency relationship holds when one or more events enable a subsequent event. The term "contingency" is chosen to indicate a sense of enablement in which human action draws upon but is not necessarily determined by elements of the environment. Contingencies are represented in a graph as arcs between events (vertices).[1] Each arc points backward in time from a single origin to one or more destinations. For example, in Figure 1, event $E3$ is contingent on event $E1$, and $E4$ is contingent on events $E1$ and $E2$.

[1]More precisely, in graph theoretic terms a contingency graph is a special kind of acyclic directed hypergraph in which contingencies are represented as hyperarcs between events represented as vertices. A hypergraph is a graph in which edges connect sets of vertices. In a directed graph, the edges are directed from one set of vertices to another and are called *arcs*. Contingency graphs are directed to represent which event is contingent upon which, and acyclic because an event can only be contingent upon prior events. Contingency graphs are specialized kinds of hypergraphs because they are restricted to arcs of form $(e_u, \{e_1, \ldots, e_n\})$ where e_u is contingenc upon $\{e_1, \ldots, e_n\}$. For example, the graph of Figure 1 includes arcs ($E3$, $\{E1\}$) and ($E4$, $\{E1, E2\}$).

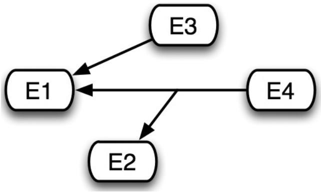

Figure 1. Contingency relationships.

Over the past two years, we have done considerable work exploring types of contingency relationships between events (Suthers, 2006; Suthers, Dwyer, Medina, & Vatrapu, 2007, 2009), including media dependencies, temporal and spatial proximity, representational similarity, and semantic overlap. These are discussed below.

Contingencies are identified through similarities in events. The most straightforward similarity between events to detect is when two events involve the same media entity and the later event depends on the state of the media entity left by the prior event. We call this kind of contingency a *media dependency*. For example, the events of opening and replying to a message are dependent on the event of creating the replied-to message (e.g. a user previously posts a message to a discussion forum), and the event of linking to or annotating an information node depends on its availability (e.g. concept map diagramming). In typical sociotechnical environments artifacts are produced, perceived, and acted upon during interaction. The sequence of these actions within and through these media entities over time form patterns of activity that are traceable partly through analysis of how such artifacts are appropriated during interaction.

Another indication of contingency occurs in synchronous interaction where *temporal proximity* implies relevance, such as in the typical reply structure of conversation (Sacks, Schegloff, & Jefferson, 1974). In visio-graphic media people also exploit s*patial proximity* and *representational similarity* to manage interaction and express association (Dwyer & Suthers, 2006; Shipman III & McCall, 1994). For example, if a representational element is given the same appearance as other elements (e.g. same color, location, or label), we might construe the change in appearance as contingent on previous uses of those visual attributes (e.g. adding an element to a group is contingent on the group's prior existence).

Tracing *semantic overlap* is more difficult. We can partially trace ideas by tracing the artifacts that express them, but actors may "transcribe" ideas to other artifacts, such as through quoting practices. (Barcellini, Détienne, Burkhardt, & Sack, 2005). More problematically for the analyst, ideas can be taken up and re-expressed in different ways.

In summary, a contingency graph is a data structure that maintains observed relations between events of interest. The graph is a representation of the evidence that has been observed for interaction between two or more actors. Contingency graphs are inclusive of the multifaceted relations that are possible (and enabled by) media-rich environments, but they need to be interpreted to identify instances of *uptake* (Suthers, 2006) in which one actor takes the results of prior activity as significant for ongoing interaction. Contingency graphs are mathematical representations of relationships in the data. They are abstract transcripts, not visualizations, but they may be visualized as diagrammatic graphs to provide notational resources for conducting analyses. Contingency graphs offer analytic affordances for discovering and understanding forms of interaction that may not be immediately apparent from micro-analysis of local episodes of interaction or from global analyses of properties of the data alone. For example, the graphs used in the study presented in this paper helped us notice a general pattern of media usage that included nonverbal as well as verbal actions. Based on this pattern we were able to more effectively frame subsequent micro analyses that relied on video data in addition to the graph structure itself.

The challenge of realizing the notational and analytic possibilities enabled by using contingency graphs prompted the development of the Uptake Graph Utility (UGU will be described in the next section). Using UGU as visual controller for example, the analyst can selectively filter elements of the graph from view, generate subgraphs based on content queries, or isolate certain structural or temporal properties of the interaction record. Figure 2 shows an example of a contingency graph taken from data presented later in this paper along with a portion of the UGU interface. As shown in the figure, a full contingency graph based solely on media dependencies and constructed from even dyadic interaction can become quite complex to interpret. The Uptake Graph Utility was developed to address this complexity by selecting those contingencies that evidence uptake. The utility is described next.

3. UGU: Uptake Graph Utility

UGU was designed and implemented to interface between log data records (stored in a relational database) and a general purpose diagramming software program (Omnigraffle™) that we used for visualizing contingency graphs (see Figure 2 above and Figure 3 below). UGU is a collection of scripts that are fused together in a panel-like desktop application. Most of the scripts communicate with the database through SQL queries. In turn, the result of each query is processed within the script in order to affect the display of the graph in the diagramming software (this requires using application specific system calls). For example, one query might be to find all events whose content has the word "disease." This query could also match events that have a contingency relationship with such "disease" events. The result of this query could then be visualized, for example, by sending a command to the diagramming tool to move all events indicated in the query result to a visible layer and hiding the rest. This illustration highlights the primary role of the utility as a tool for manipulating the visualization of a contingency graph. UGU has two general functions that we characterize as *transcript transformation* and detailed *content analysis*, described in the next two sections.

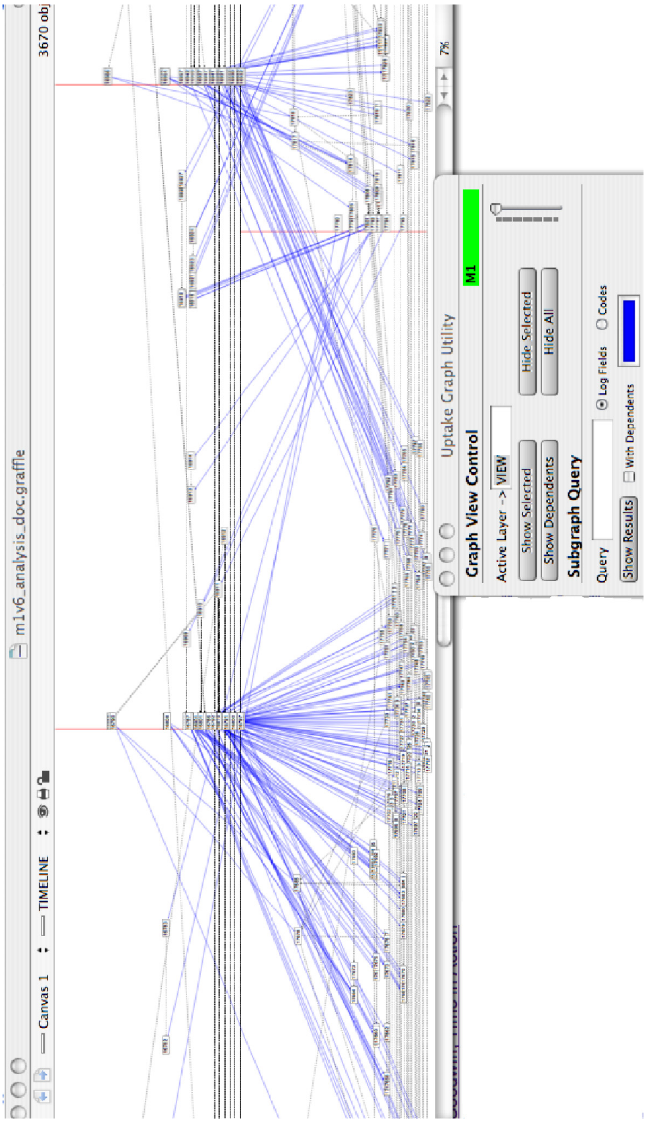

Figure 2. Portion of initial contingency graph (based on media dependencies) displayed in Omnigraffle™ above, with part of the UGU control panel visible below. See Figure 3 for the full UGU panel.

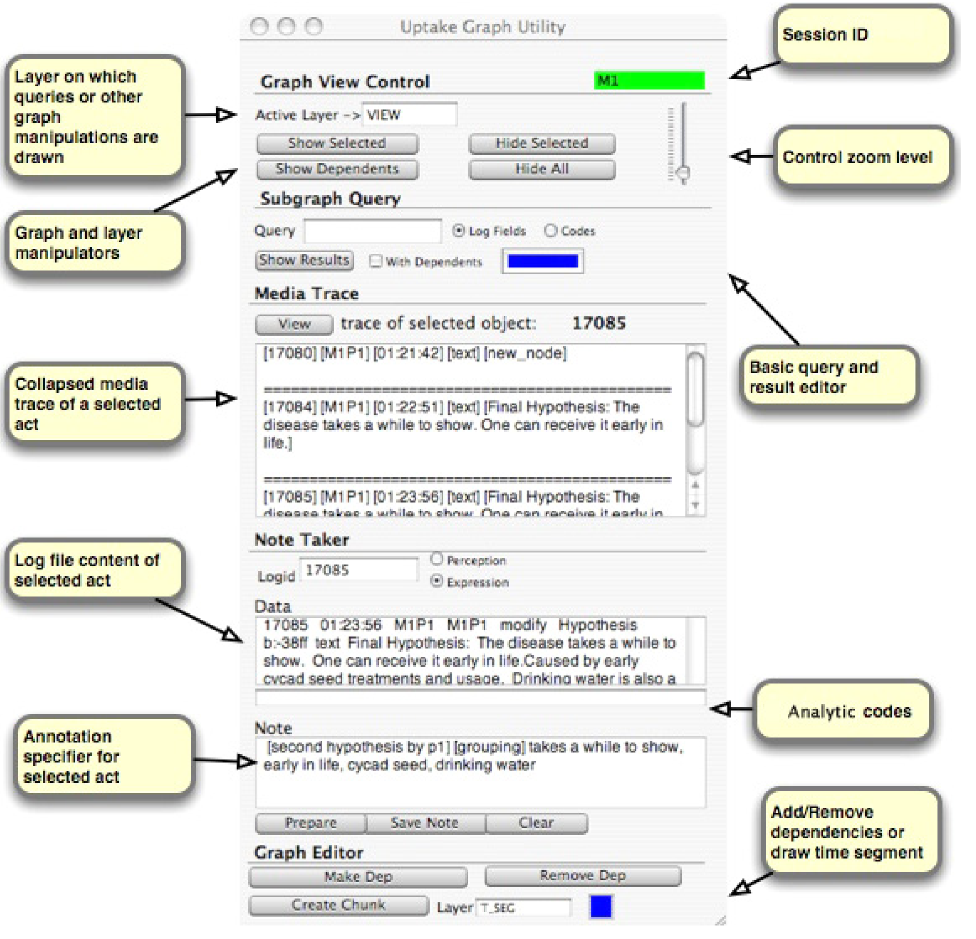

Figure 3. The Uptake Graph Utility pallete.

3.1.1. *Transcript transformations*

Certain manipulations of the data representation should support the analyst in reasoning about the data in such a way as to make some aspects of the transcript more or less salient. Transformations provide useful *constraints* for coping with the multifaceted nature of sociotechnical interaction (Sanderson & Fisher, 1994). They can include graphic manipulations of color, scale (zoom level), and visible/non-visible layering. UGU supports representational transformations through the use of the *Graph View Control*, *Subgraph Query*, and to a lesser extent, the *Graph Editor* interface components (shown in Figure 3). These components allow the analyst to re-present the contingency graph in correspondence with particular analytical questions. Transformations can be applied at the granularity of a subgraph or vertices. For example, if one wishes to trace the path through the graph (a subgraph) of a particular media object, it should be possible to illuminate that path and distinguish it from the others. This can be achieved in multiple ways. Using the *Graph View Control* as an option one can specify that the selected elements (vertices and arcs) of the graph render on the *active layer*. The *Subgraph Query* component further supports this practice by allowing the user to specify what highlight color to use for visible elements and whether or not to include the contingency arcs in the rendering of the contingency graph.

3.1.2. *Content analysis and annotation*

The second general function of the Uptake Graph Utility is to provide comprehensive access to the content of vertices (events). This is provided by the *Media Trace* and *Note Taker* components (see Figure 3). The *Media Trace* displays a listing of a sequence of contingent events, one of which is selected by the analyst (Figure 3). A discussion thread, for example, may be displayed using the media trace component by selecting the parent posting. An analytic reading can be done with regard to the thread without having to navigate or scroll through potentially wide graph spaces. The *Note Taker* component displays and allows creating of new associations between a selected vertice or arc and an analytic comment. These annotations are written to and persist as meta information in the data record.

3.1.3. *Additional analysis support*

In addition to the above, UGU also provides support for identifying new contingencies during the course of analysis. The *Graph Editor* component supports this as well as the practice of segmenting the sequential data based on analytically determined boundaries. For example, regions of the graph can be specified to correspond with episodes or interaction sessions.

Figure 4 illustrates an overview of our analytical process beginning with computer log files. As will be shown in the case study, the contingency graph is useful for discerning patterns of interaction and providing indications for further investigation into other data sources. The next section describes our analytic method followed by presentation of the case study.

4. Methods

4.1. *Data collection and contingency graphs*

The vertices of contingency graphs should be based on actual events recorded in raw data, including video, video transcriptions, or log files. In general, any data that can be parsed into distinct acts that constitute evidence of actual events is valid input. In our case, we used an XML representation of log files generated by a peer-to-peer collaborative learning environment as our initial data source (bottom of Figure 4). The logs represent interactions between two participants using the software.

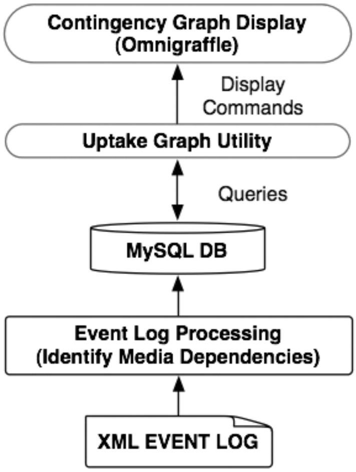

Figure 4. Process overview.

Our initial step in preparing the data for analysis is to construct a contingency graph based on media dependencies (Sec. 2.1.2), as shown in Figure 4. We call such a graph a *media trace*. It is convenient to begin with this type of dependency because our log file events contain a field that records a media object identification string. Each media object (e.g. graph node, discussion posting, etc.) created using the software during the interaction is assigned a unique identifier. Once a media object is created, its assigned identifier is included in all subsequent user log events that implicate the object. For example, if a user replies to a prior discussion posting, the log event will record the details of the new message and will include a reference to the prior message via its unique media identifier.

Table 1 below shows a snippet of our log file in which participant $P1$ creates a graph node containing a message about aluminum. The next log event indicates that the other participant $P2$ has received the information node (on their machine), has accessed it, and changed its location on the screen. The fourth event in the sequence indicates that $P1$ has edited the original node by adding additional text. One property to note from Table 1 is the inclusion of nonverbal events in the action sequence. "Move" events, for example, include which object was moved and where it was moved to (screen coordinates are not shown in the example).

Table 1. An example media trace based on media dependencies.

Localtime	Creator	Object id	Action	Content
0:05:13	M8P1	7e0d	modify	It could be the aluminum in the drinking water.
0:13:56	M8P2	7e0d	modify	peerHasRead
0:13:56	M8P2	7e0d	move	
0:15:48	M8P1	7e0d	modify	It could be the aluminum or something in the drinking water.

Table 2. A sample result set of a media trace using the object id from one participant's machine log where the shaded area signifies a subsequence of update events from another participant (This trace is also visualized in Figure 5).

Logmsgid (P1)	Logmsgid (P2)	Localtime	Creator	Object id	Media type	Action	Content
16463	17424	0:29:48	P1	-6c24	Hypothesis	modify	Disease caused by aluminum.
16589	17526	0:39:06	P1	-6c24	Hypothesis	move	
16606	17459	0:41:07	P2	-6c24	Hypothesis	modify	peerHasRead
16618	17474	0:41:07	P2	-6c24	Hypothesis	move	
16658	17514	0:41:07	P2	-6c24	Hypothesis	move	
16678	17625	0:47:04	P1	-6c24	Hypothesis	move	

The sequence of events listed in Table 1 illustrate how media dependencies can be identified using information stored in the log entries. In this case, the events share a common object identifier (7e0d) that can be delineated with respect to which user performed the act and when the action was performed. We leveraged this property of the log files to design a program that builds an initial contingency graph of media traces.

Media traces are stored as a table in the database containing the log (middle of Fig. 4). A single trace is a set of event pairs where each pair indicates a *source* and *destination* identifier. Table 2 above lists an example trace for object id **-6c24** and user **P1**. The shaded area in the table indicates "received" log events from the other user **P2** (for the ensuing discussion it is important to note that any *logmsgid* value uniquely identifies an event in the log database). The sequence of update events (shaded rows in Table 2) and the local event that follows it (*logmsgid* 16678) form a one to many relationship. That is, the act at 16678 is contingent on the *set* of prior acts assembled as a series of updates from the partner's machine (16606 thru 16658).

The final step in preparing the contingency graph for analysis is transcoding the database table generated by the media trace discussed above to an XML file format compatible with the diagramming software used for external rendering of the graph. Figure 5 illustrates the visualization of contingencies derived from the media trace segment detailed in Table 2.

Figure 5 shows how a careful interweaving of separate event sequences (one from each user) into a unified sequence can capture empirical evidence for relations between interactors. Events on $P1$'s workstation are above the timeline and events on $P2$'s workstation are below the timeline. The arcs that intersect the timeline connect event actions between the actors. It is important to note that our log files contain nonverbal actions such as moving graphical objects on the screen. These "move" events are also represented in the contingency graph. The multiple arcs emanating from vertice 16678 is evidence that multiple actions were initiated by participant 2 prior to participant 1's workspace update. The multiple contingencies demonstrates evidence that participant 1's next act is contingent on potentially multiple actions by participant 2.

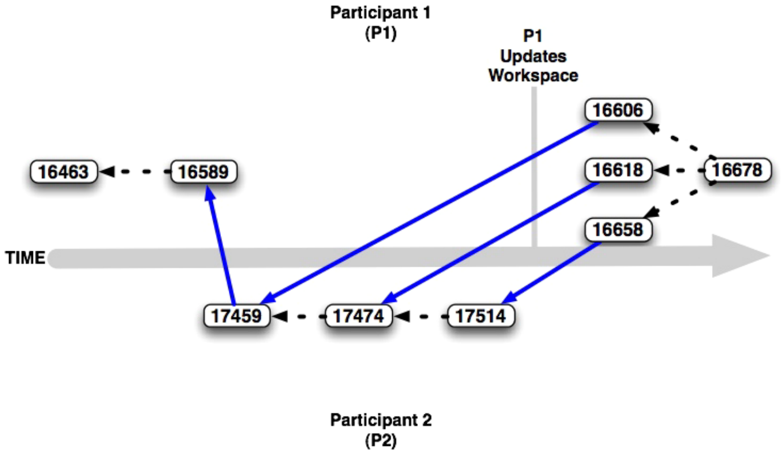

Figure 5. Sample segment of contingency graph based on media dependencies for object listed in Table 2.

In summary, our initial contingency graph is based on media dependencies as these kind of relationships can be readily mined from the log file data. Once our graph is prepared (stored and visualized) we employ two general analytic practices, *segmentation* and *tracing*, for working with the graph during analysis.

4.2. *A general analytic approach*

Segmentation (or chunking) is a way to shape the data into discrete partitions or episodes that are the basic units of analysis (Jordan & Henderson, 1995). Segmentation can operate at multiple granularities. For example, in the case study given later in this paper, we began with individual acts of media manipulations and contingency relationships between them as the initial units of analysis, but then chunked subgraphs of contingencies into episodes of recognizable activity evidenced by participant's orientation to such activity (Garfinkel, 1967).

Tracing is the act of identifying certain elements of interest as reference points, and then moving forward or backward along pathways in the graph to unravel the interaction trajectory in which those elements were formed. The rationale for following a path is analytically motivated. At any given moment in the analysis, new entities uncovered may or may not warrant the definition of another data chunk, may inspire subsequent traces, may induce a closer examination of the data sources such as video, or require establishment of new contingencies previously unidentified. The analytic strategy taken in this paper follows that of our previous work (Suthers *et al.*, 2007, 2009). In that work, we looked at post-interaction essays composed individually by participants to determine what each person concluded about the problem posed to them. We then traced back from these acts of writing through contingency relationships to identify interactional sequences of acts that could potentially account for the participants' conclusions.

In summary, the strategy taken for the current analysis was to render the graph such that one can identify relationships between vertices and their position within extended trajectories. To limit the amount of screen clutter, only the log id for each act is made visible on each vertex. In order to allow ready-at-hand access to the data portion of each act, the contents of each event are written into the "Note" facility available in the diagramming software application. On mouseover, the application provides a pop up window displaying the contents of that element (notes are also visible in a floating palette). Other functions of the diagramming software were marshaled for the purpose of trying to work with a contingency graph as an analytical artifact. These included the use of layers, zoom, color, and sophisticated object selection mechanisms (e.g. selecting graphical elements with similar visual properties). After experimenting with the contingency graph using these built in tools it became clear that more control was needed with respect to how the application rendered aspects of the graph. In particular, a coupling to the database was desired in order to support analysis driven queries in an efficient and systematic manner and to integrate those results with the visualization. This requirement and the emergence of our own approach to analysis spurrred the development of the Uptake Graph Utility as a way to support these analytical practices. We now turn to an analysis that makes use of the contingency graph, UGU functions, and video data to uncover and understand the negotiation of representational practices in a collaborative learning environment.

5. Case Study

The following case study illustrates a pattern of interaction between two individuals engaged in a joint-problem solving exercise while using a shared networked workspace environment (Figure 6). The data is drawn from an experimental study conducted for purposes reported in detail elsewhere (Suthers *et al.*, 2008).

Dyads were recruited from introductory university natural science courses. Using informational materials we provided in the workspace (e.g. upper left of Figure 6), they worked to identify possible causes of a disease in Guam, ALS-PD (Amyotrophic Lateral Sclerosis-Parkinsonism Dementia complex). The software used by these participants provided both threaded discussion (lower left of Figure 6) and graphical evidence mapping tools (right side of Figure 6) derived from Belvedere (Suthers *et al.*, 2001). The session took place over the course of approximately two hours. Participants were in different rooms, and each participant's view of the shared environment was updated using a software protocol that distributed respective workspace changes at intermittent times during the interaction, thereby simulating asynchronous interaction typical of online learning (see Suthers *et al.*, 2008 for full rationale and discussion). Participants would occasionally "take a break" from the problem to play a game of Tetris$^{\text{TM}}$. When they returned to the workspace, changes from their partner since the last break were displayed step by step.

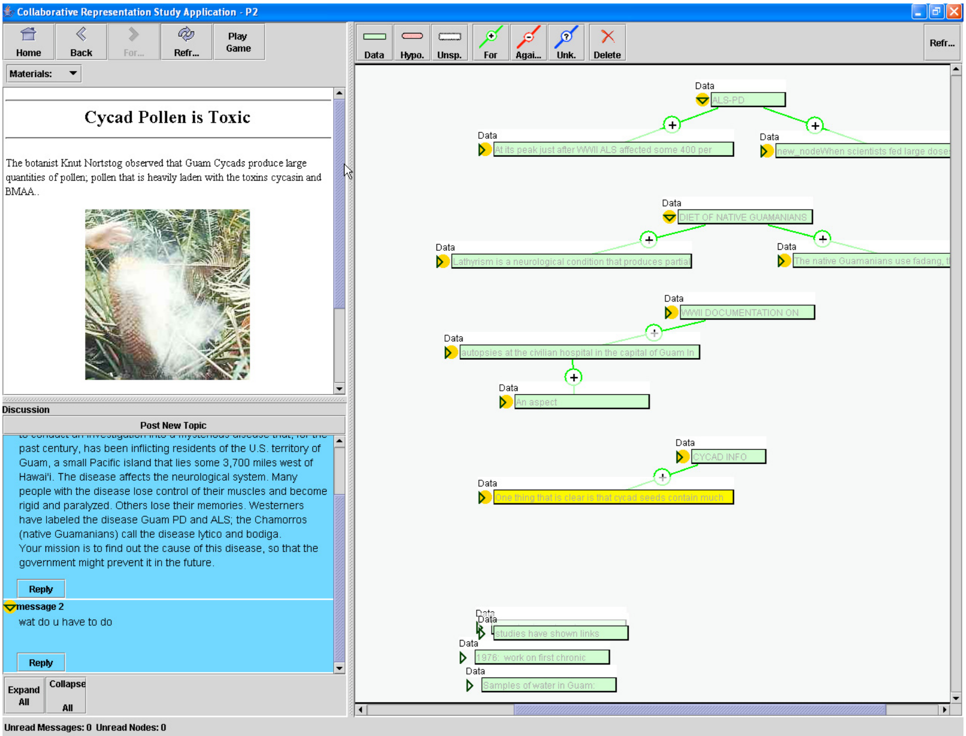

Figure 6. Screen display of one participant's view of the shared workspace.

The original study used statistical methods to test hypotheses concerning the effects of different software designs. Follow-up analyses of interaction were undertaken to understand how participants appropriated the media in their collaboration. The analysis reported in this paper specifically sought to account for ways in which one dyad both converged and diverged in their interpretations of causes of ALS-PD, by tracing out sequential patterns of representational practices (Kozma, 2003; Roth, 2003). The analysis highlights an evolving transformation of a collaborative representational practice. These practices and the artifacts left in their wake provide an explanation for the conceptual convergence and divergence in the conclusions expressed by each participant.

5.1. *Episode: concluding work*

The analysis begins with an important reference point in the interaction, a sequence of activity in which both participants, labeled $P1$ and $P2$, express conclusions concerning the possible causes of ALS-PD. This episode takes place in a time span of approximately 10 minutes towards the end of the session. The beginning of this episode is indicated by $P1$'s prompting for a conclusion ([17028] in contingency graph of Figure 7 below). $P1$ makes this request using a discussion posting. Despite

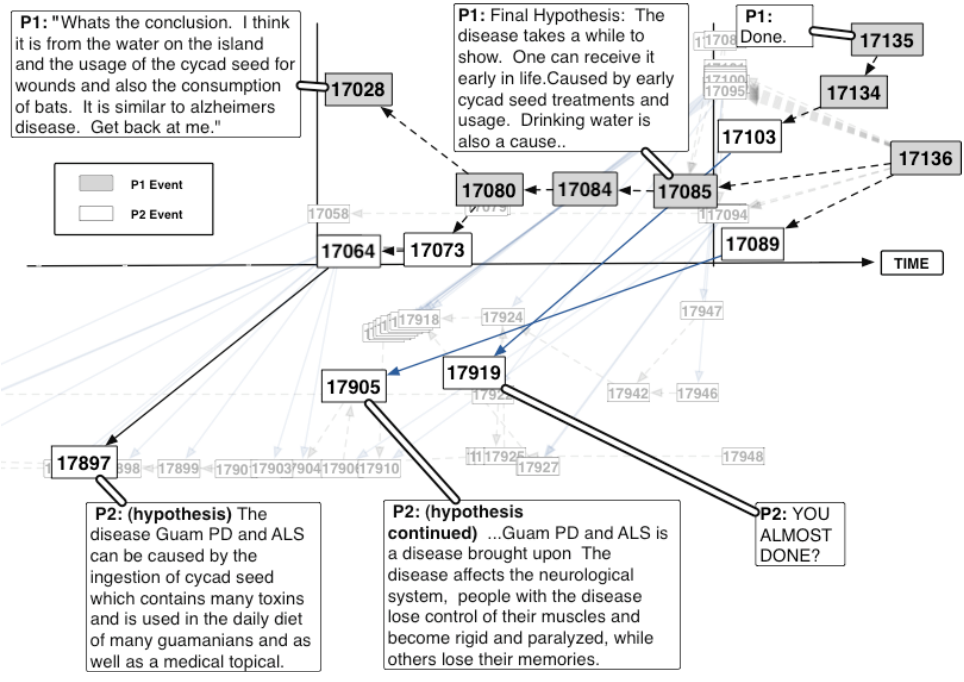

Figure 7. Subset of contingency graph showing the conclusions expressed by both participants. Large boxes display the content manipulated in selected events.

the fact that $P2$ does not read the message ($P2$ did not initiate further requests for workspace updates), it indicates $P1$'s plans to initiate a negotiated conclusion. This is evidenced by a subsequent act that proposes a "final" conclusion [17085] incorporating formative elements of $P2$'s concluding work, [17897], which coincidentally begins and is concurrently developed by $P2$ [17905], at approximately the same time as $P1$'s (unread) request. The episode ending is negotiated when $P2$ asks whether $P1$ is done [17919]. $P1$ reads and responds by stating; "Done" [17135] then immediately makes a final "For" link between $P1$'s and $P2$'s hypothesis node [17136].

Each person integrated information shared during their prior work into their respective fully developed conclusions. We are able to trace information sharing because we used a "hidden profile design" (Stasser, 1992) in which information was originally distributed across participants such that it must be shared to reach an optimal conclusion. "Cycad usage" is cited in each person's final conclusion. This is significant because both participants consider multiple specific hypotheses but only converge on one.

Given that very little linguistically explicit negotiation concerning hypotheses took place during the interaction, an analytic trace was initiated to provide an interactional account of this convergence on "Cycad usage". Also, within this segment

$P1$ makes “drinking water” a salient factor in her hypothesis, [17028 & 17085], while $P2$ does not cite drinking water at all, although information referencing “drinking water” was previously shared during interaction. Another analytic trace was generated to account for this divergence. The challenge presented by this pair’s interaction is that it involved numerous evidence map manipulations. Video screen capture of the session was used to assist in interpreting their actions. The contingency graph served as an index into the video data as needed.

Details of our analysis are presented in the following sections to show how our analytic practices were supported by the Uptake Graph Utility. We present the trace concerning “drinking water” first because it uncovered a larger scale phenomenon that led to our understanding of the significance of inscriptional manipulations with regard to the agreement on cycad convergence. Thus, this sequence of examples illustrates how analysis changes granularities, from tracing out relations between individual acts of media manipulations to relations between episodes of such manipulations and back to fine-grained analysis of appropriation of graphical resources.

5.2. *Trace of conceptual divergence uncovers representational practices*

One of the consistent concepts indicated in $P1$’s argument during their concluding work is that “drinking water” is one possible cause for the disease. An attempt to build an account of this concept through the interaction history began by forming a query (input into UGU) in order to highlight acts that reference that text string and the contingencies between those acts. The highlighted vertices in the resulting contingency graph revealed references to “drinking water” that were included in the information provided to $P1$ in relation to aluminum as a potential cause of the disease. A second query was invoked to capture acts that also referenced aluminum, extending the trace. The resulting contingency graph is summarized schematically in Figure 8. We have annotated the contingency graph to summarize some of the content expressed or actions taken. In two particular instances $P1$ shares information with $P2$ related to the contamination of drinking water by aluminum. $P2$ *acts upon* this information by performing a series of moves evidenced by clumps of move events in the graph. These acts by $P2$ do not contain linguistic responses; only a series of moves (drag and drop acts) performed within the evidence map. This pattern is consistent throughout the remaining portions of the session. The trace shown in Figure 8 could indicate that $P2$ is moving nodes around in order to see them, or to get them out of the way: dragging and dropping of graphical objects for these reasons is frequent. In this case however, the periodic-like pattern of $P2$’s series of movements induced us to explore the video record for these episodes.

The video shows that $P2$ is not randomly moving nodes around, but performing a series of evidence map *reconfigurations* to organize information previously shared during the session. After $P1$ contributes new information, $P2$ moves nodes to create

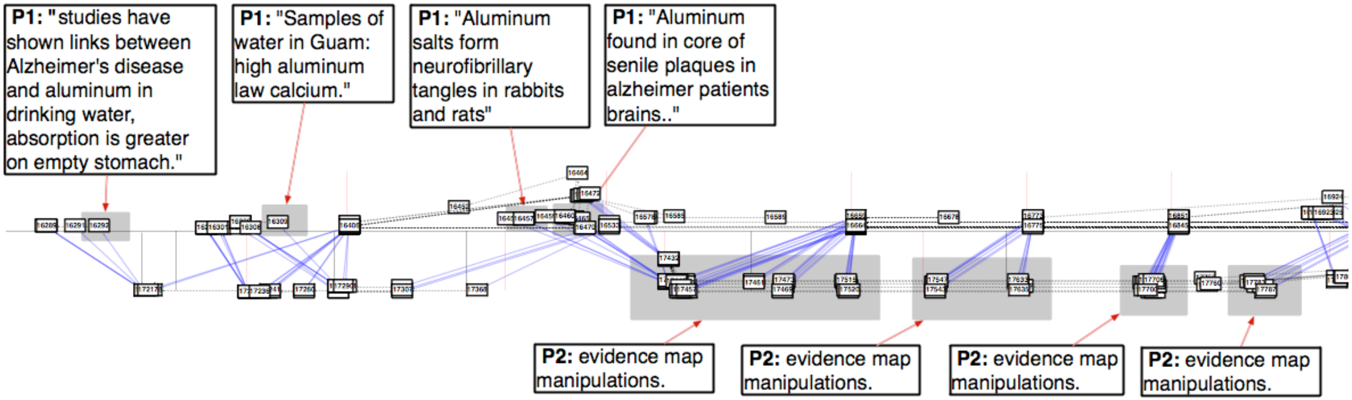

Figure 8. Information sharing by $P1$ followed by systematic graph manipulations by $P2$ during the first hour of the session.

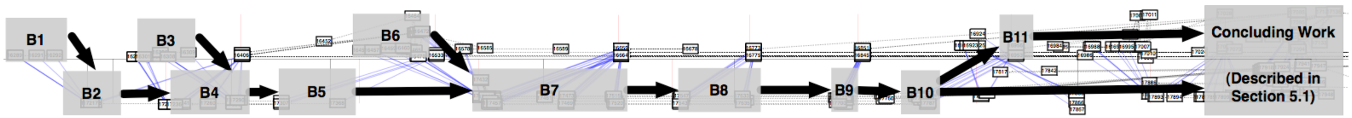

Figure 9. High level (episodic) view of interaction over the entire session.

spatially distinct groups that provide conceptual delineation. In addition to this spatial organization, both participants create links between nodes within groups that further clarify their inclusion in the group. (Their work will be illustrated in detail in the next section.)

Figure 9 illustrates this same trace at a higher level of abstraction, as a series of related episodic segments. Beginning at the left, $P1$ shares information containing a reference to aluminum in water as a contaminant in the first two segments $[B1, B3]$. The third information-sharing event by $P1$ contains two references that correlate aluminum and neurological symptoms of ALS-PD $[B6]$. The reaction to the three sharing acts by $P2$ is shown as series of evidence map manipulations $[B2, B4, B5$ & $B7$-$10]$. Intersubjective uptake is indicated by $P2$'s visual transformation of the shared information nodes and is followed by a series of intrasubjective transformative acts on the part of $P2$, who continually appropriates the relation-indicating power of the evidence map. The fact that there is very little related action on the part of $P1$ during these episodes indicates that $P2$ is accountable for subsequent transformations. As shown on the far right of the diagram, intersubjective acts again occur as the concluding work segment discussed above, is initiated (the right side of Figure 9 corresponds to Figure 7).

Participants' grouping acts form a representational configuration that foreshadows each participant's concluding work. $P2$'s creates two separate groupings, among others, for the information containing aluminum: as an agent in metal intoxication and as water contaminant. One explanation for the divergence on this concept is that the resulting visual organization provides a selection context from which each participant performs his or her concluding work. The emergent representational artifact, the evidence map, facilitated multiple meanings for each participant to appropriate in conceptually convergent and divergent ways. $P1$ apparently appropriates this representational scheme initially enacted by $P2$ with a slightly divergent interpretation.

5.3. *Episode: appropriation of representational practice*

We turn now to a closer look at $P2$'s organization of the graph, and $P1$'s appropriation of this practice in the handling of data about cycads as a potential disease agent. The contingency graph enabled us to provide a representational practices account of their conceptual convergence. When visualized, the contingency graph exposed patterns of interaction; provided direct pointers, via timestamps, to relevant locations in the video record; and provided frames of reference for interpreting the video. This framing made the interrelation between the two separate video streams salient for determining the emergence of a shared representational practice.

At the beginning of the session, $P2$ creates and organizes data nodes into conceptual groups. These groupings are specified through spatial proximity and the use of links between nodes. Figure 10 is a screenshot of $P2$'s screen after having constructed such an initial graph configuration. An important dimension of these

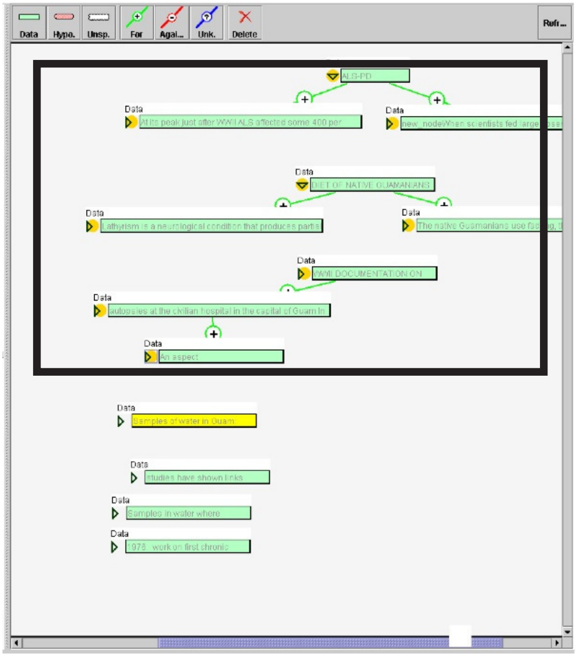

Figure 10. $P2$'s screen, initial configuration.

grouping configurations is that topic-based nodes are positioned as hubs to conceptually related information nodes.

At the bottom left of the graph workspace in Figure 10 are four nodes that were rendered on $P2$'s screen as a result of a recent update from $P1$. At this point we see two distinct representational practices. $P2$ has made conceptual organization a dominant method for representation. $P1$'s groupings show a less defined representational practice. Later, having received a series of updates of $P2$'s organizational work, $P1$ demonstrates appropriation of $P2$'s practice. Figure 11 shows the state of $P1$'s screen after she has created additional nodes and grouped them into a visual configuration that resembles $P2$'s scheme. In addition to visual organization, $P2$'s conceptual structure is adopted, as $P1$ orients the nodes towards a central conceptually labeled node. This node also represents an explicit expression of a hypothesis, "Disease caused by aluminum" which reveals $P1$'s practice of articulating hypotheses through language (not adopted by $P2$).

A concurrent activity during the episode depicted in Figure 11 occurs on $P2$'s machine. Figure 12 shows the introduction of cycad information into the evidence map. Following his own representational convention, $P2$ positions the label, "cycad info", and three related data nodes into an identical configuration as the others. Figure 13 shows a subsequent act on $P1$'s machine where she introduces a data node containing information about "cycad". (Time has elapsed, so $P1$'s screen reflects

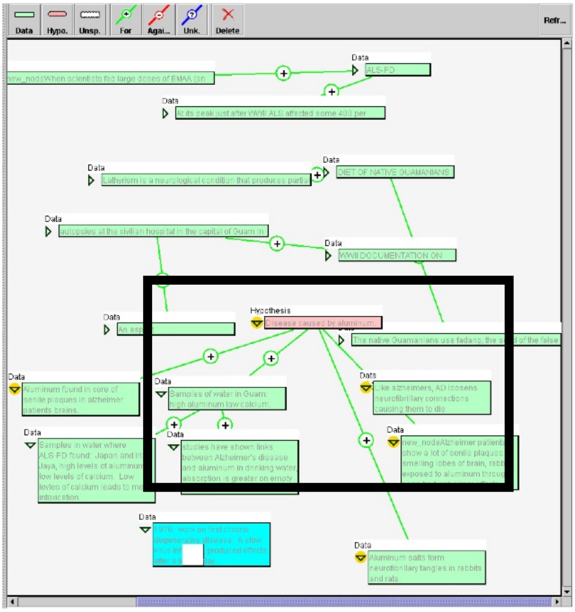

Figure 11. $P1$'s screen, appropriation of representational grouping practice.

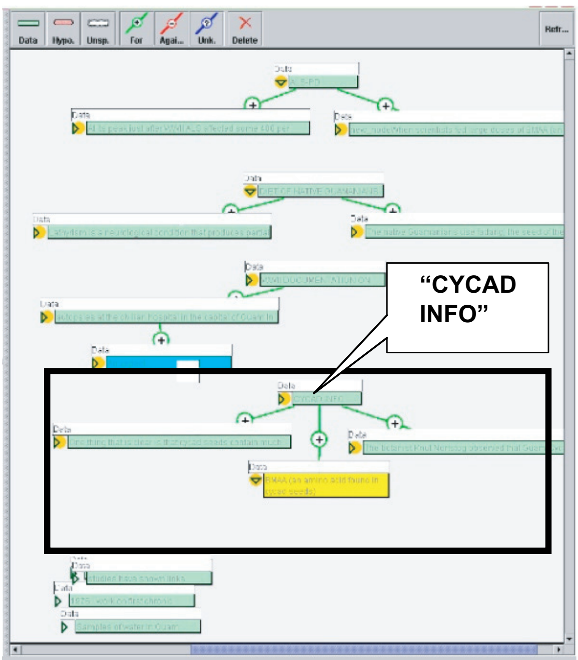

Figure 12. $P2$ creates “cycad” grouping.

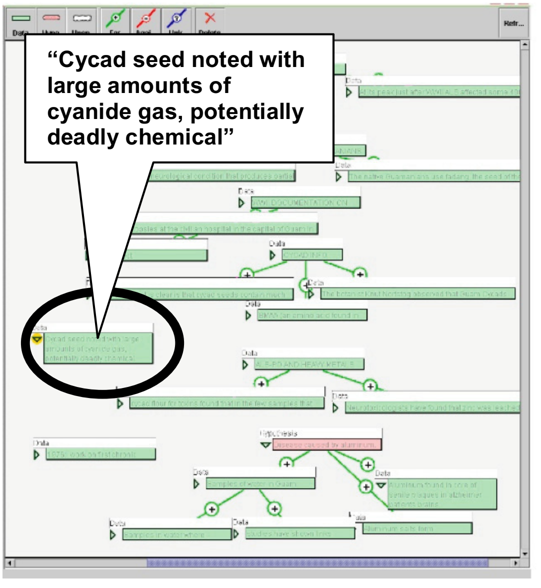

Figure 13. $P1$ creates a "cycad" data node.

the ongoing work of the two participants.) In this context, a "cycad" related node is created and positioned in a somewhat arbitrary location with regard to the ongoing visual grouping. On receiving an update from $P1$ containing the cycad data, $P2$ reads the contents of the node, then drags the node to an inclusive position of the cycad conceptual grouping (circled in Figure 14). $P2$ then follows this repositioning with the creation of a link between the node and the "CYCAD INFO" hub, further expressing its group membership.

Subsequently, each participant brings "cycad usage" forward in distinct ways. $P1$ articulates cycad salience through a statement placed in a hypothesis node, "Disease caused by cycad seed usage" (Figure 15, left side), while $P2$ posts a short "themed" node expressing "USES OF CYCAD" (Figure 15, right side). Each participant without knowledge of the other performs these respective acts. They coincidentally indicate cycad usage at approximately the same time. In addition to posting her hypothesis node, $P1$ integrates it into the "CYCAD INFO" group configuration by creating four links to supporting data. $P2$ also groups and links data nodes to their expression ("USES OF CYCAD"). It is a mutual appropriation of a grouping practice. $P1$ and $P2$ both begin wrapping up their work within five minutes after this episode and thus initiate the concluding work episode presented above (Figure 7).

6. Discussion

The above analysis provides an explanation of one aspect of how the two participants converged and did not converge on conclusions in a joint problem-solving

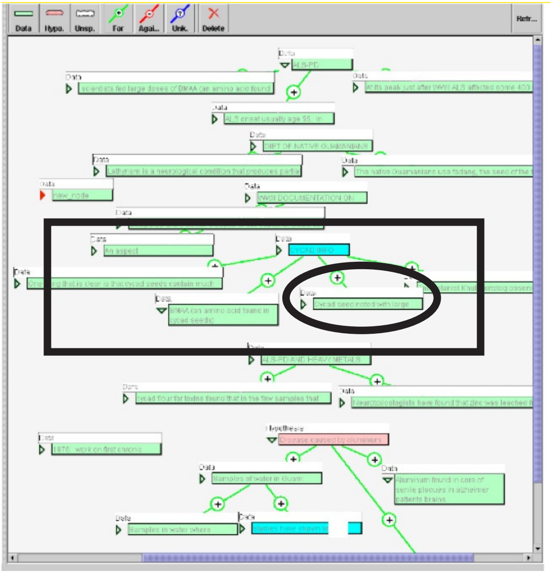

Figure 14. $P2$ receives "cycad" data node from $P1$ (Figure 13) and repositions and links it into a cycad group (Figure 12).

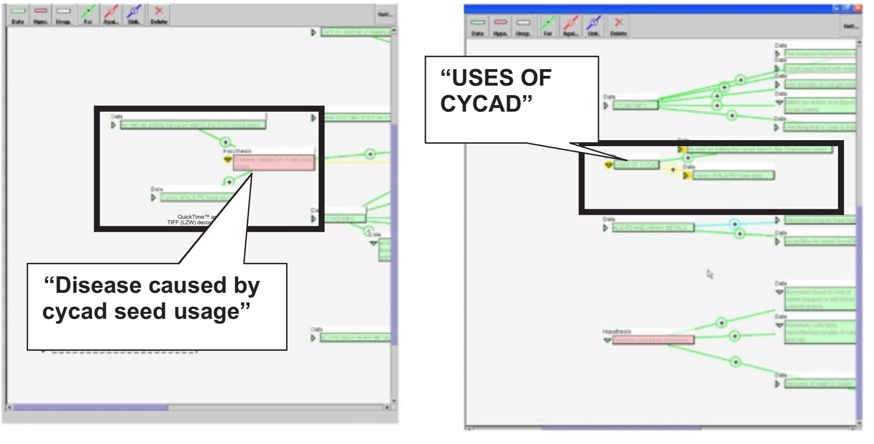

Figure 15. $P1$ and $P2$ articulate new cycad groupings independently.

task. The goal was to explore an application of the uptake analysis framework, its representation and praxis, by way of taking a detailed look at how interactions through shared representations might account for particular instances of conceptual convergence and divergence. We uncovered a case of interactional negotiation of representational practices, implicitly proposed by one participant ($P2$) through demonstration, and taken up by the other participant. This shared representational practice was particularly apparent in their handling of the question of cycad seeds as a potential cause of the disease in question, a finding that is consistent with the fact that this was the one cause on which both participants agreed. In contrast, information about the role of drinking water in relation to aluminum was distributed in the graph in a manner that seems consistent with the lack of agreement on the importance of this element and the fact that $P2$ received the bulk of evidence against the associated aluminum hypothesis.

The primary significance of this analysis is that negotiated representational practices can be found in asynchronous interactional settings, and can influence outcomes of a collaborative session. Such phenomena merit further study to understand how learning is accomplished (Koschmann *et al.*, 2005). However, due to the fact that interaction is not immediately salient in asynchronous online settings, representations and tools that make interaction patterns visible are needed. We have prototyped one such tool, and propose to develop further tools based on the contingency graph, transcending differences in log file formats and the distribution of interaction across media.

7. Conclusion

Following Reimann (2009), we believe that "time is precious": the practices of participants interacting in technology mediated environments are best understood by tracing how they unfold as processes over time, rather than by analytic methods that abstract completely away from time. The historical trajectory of joint action captured in discrete software events underscores sequential structures and patterns that are difficult if not impossible to grasp through traditional transcriptional "readings" of the data. Many researchers, for example, study online interaction by focusing on reply structure (e.g. Hewitt, 2003; Sack, 2001). Although reply structure is a good place to start, offering clear evidence of uptake, our method acknowledges the many other ways in which one person's action may build on those of others. This picture is further complicated by the distribution and contingency of these actions across people and media.

One approach taken by us and others is to recognize a more general relation between units of action as the starting point for uncovering more elaborate relational structures or graphs. Lonchamp (2009), for example, proposes a method for identifiying "generalized conversations" for tracing dialogic interaction that supports subsequent analyses such as knowledge building. This bottom up approach is also seen in other work that attempts to situate analysis of learning environments and learning upon grounded accounts of joint action. Barab, Hay & Yamagata-Lynch's (2001) Construction of Networks of Action Relevant Episodes perhaps has the greatest affinity to our approach. Their method of selecting "tracers" is akin to our tracing of media object IDs or of concepts back through the contingency graphs. We share, with these researchers, a view that a representation of the contingent nature of interaction is a foundational requirement for analysis of phenomena at larger granularities. Barab *et al.* (2001) refer to this as a process of unification in which multiple time scales and contexts are captured in one analysis. In our case, the use of contingency graph visualizations shows how analysis can be scaleable, relying on micro and episodic sequences in the accounting of computer-supported collaborative learning.

The contingency graph is an abstraction of what is traditionally thought of as a transcript. Enabling its representation and its relational structure in a computationally accessible format promises to support sophisticated and scaleable analytical practices. Segmentation and tracing are two such practices that are fundamental in working with relationally represented sequential data. It is through cycles of segmentation and tracing that one is able to isolate aspects of interaction under investigation. One danger in isolating data elements in this way is the potential for decontextualization — losing sight of the full breadth of contingencies at play in a given context. To counter this tendency, we employed video analysis to complement our approach. In practice, the visual distinction between selected and non-selected acts can be leveraged within cycles of selective transformation. The Uptake Graph Utility was developed around these two ideas and provides a prototype for planned further development of related tools based on the contingency graph representation.

Acknowledgments

The original study on which this work was based was supported by the National Science Foundation under award 0093505.

References

Aviv, R., Erlich, Z., Ravid, G., & Geva, A. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal of Asynchronous Learning Networks*, *7*(3), http://www.sloan-c.org/publications/jaln/v7n3/v7n3_aviv.asp.

Barab, S. A., Hay, K. E., & Yamagata-Lynch, L. C. (2001). Constructing networks of action-relevant episodes: An *in situ* research methodology. *The Journal of the Learning Sciences*, *10*(1&2), 63–112.

Barcellini, F., Détienne, F., Burkhardt, J.-M., & Sack, W. (2005). A study of online discussions in an Open-Source Software Community: Reconstructing thematic coherence and argumentation from quotation practices. In *Proceedings of the Communities and Technologies Conference*. Milan, Italy.

Bruckman, A. S. (2006). Analysis of log file data to understand behavior and learning in an online community. In J. Weiss, J. Nolan, J. Hunsinger, & P. Trifonas (Eds.), *International Handbook of Virtual Learning Environments* (pp. 1449–1465). New York: Springer.

Danish, J. A. & Enyedy, N. (2006). Negotiated representational mediators: How young children decide what to include in their science representations. *Science Education*, *91*(1), 1–35.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, *46*(1), 6–28.

Dwyer, N. & Suthers, D. D. (2005). A study of the foundations of artifact-mediated collaboration. In T. Koschmann, D. D. Suthers, & T.-W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years!* (pp. 135–144). Mahwah, NJ: Lawrence Erlbaum Associates.

Garfinkel, H. (1967). *Studies in Ethnomethodology.* Englewood Cliffs, New Jersey: Prentice-Hall.

Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal of Educational Computing Research*, *28*(1), 31–45.

Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers & Education*, *41*, 397–420.

Jordan, B. & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, *4*(1), 39–103.

Koschmann, T., Zemel, A., Conlee-Stevens, M., Young, N., Robbs, J., & Barnhart, A. (2005). How do people learn: Member's methods and communicative mediation. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers and biases in computer-mediated knowledge communication (and how they may be overcome)* (pp. 265–294). Amsterdam: Kluwer Academic Press.

Kozma, R. (2003). Material and social affordances of multiple representations for science understanding. *Learning and Instruction*, *13*(2), 205–226.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2), 259–284.

Larusson, J. A., & Alterman, R. (2007). Tracking online collaborative work as representational practice: Analysis and tool. In C. Steinfield, B. Pentland, M. Ackerman, & N. Contractor (Eds.), *Communities and Technologies 2007: Proceedings of the Third Communities and Technologies Conference*, Michigan State University (pp. 245–264). London: Springer-Verlag.

Lonchamp, J. (2008). A three-level analysis of collaborative learning in dual-interaction spaces. *International Journal of Computer-Supported Collaborative Learning*, *4*(3), 289–317.

Martinez, A., Dimitriadis, Y., Rubia, B., Gomez, E., & de la Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & Education*, *41*, 353–368.

Overdijk, M. & van Diggelen, W. (2008). Appropriation of a shared workspace: Organizing principles and their application. *International Journal of Computer-Supported Collaborative Learning*, *3*(2), 165–192.

Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *Computer Supported Collaborative Learning*, *4*(3), 239–257.

Roschelle, J. (1996). Designing for cognitive communication: Epistemic fidelity or mediating collaborating inquiry. In D. L. Day & D. K. Kovacs (Eds.), *Computers, Communication & Mental Models* (pp. 13–25). London: Taylor & Francis.

Roth, W.-M. (2003). *Toward an anthropology of graphing: Semiotic and activity theoretic perspectives.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Sack, W. (2001). Conversation map: An interface for very large-scale conversations. *Journal of Management Information Systems*, *17*(3), 73–92.

Sanderson, P. & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human-Computer Interaction*, *9*(3–4), 251–318.

Shipman, F. M., III & McCall, R. (1994). Supporting knowledge-base evolution with incremental formalization. Paper presented at the *ACM CHI 94 Human Factors in Computing Systems Conference*. From http://www.acm.org/pubs/articles/proceedings/chi/191666/p285-shipman/p285-shipman.pdf.

Stahl, G. (2006). *Group Cognition: Computer support for collaborative knowledge building.* Cambridge, MA: MIT Press.

Stasser, G. (1992). Information salience and the discovery of hidden profiles by decision-making groups: A "thought experiment". *Organizational Behavior and Human Decision Processes*, *52*(1), 156–181.

Suthers, D. D. (2006). A qualitative analysis of collaborative knowledge construction through shared representations. *Research and Practice in Technology Enhanced Learning*, *1*(2), 1–28.

Suthers, D. D., Connelly, J., Lesgold, A. M., Paolucci, M., Toth, E. E., Toth, J., & Weiner, A. (2001). Representational and advisory guidance for students learning scientific inquiry. In K. D. Forbus & P. J. Feltovich (Eds.), *Smart Machines in Education: the coming revolution in educational technology* (pp. 7–35). Cambridge, Massachusetts: AAAI Press, The MIT Press.

Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2007). A framework for eclectic analysis of collaborative interaction. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *The Computer Supported Collaborative Learning (CSCL) Conference 2007* (pp. 694–703). New Brunswick: International Society of the Learning Sciences.

Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2009). Exposing interactional processes in online learning. In K. Kumpulainen, C. E. Hmelo-Silver & M. César (Eds.), *Investigating classroom interaction: methodologies in action* (pp. 121–144). Rotterdam: Sense Publishers.

Suthers, D. D. & Hundhausen, C. (2003). An experimental study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences*, *12*(2), 183–219.

Suthers, D. D., Vatrapu, R., Medina, R., Joseph, S. & Dwyer, N. (2007). Conceptual representations enhance knowledge construction in asynchronous collaboration. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *The Computer Supported Collaborative Learning (CSCL) Conference 2007* (pp. 704–713). New Brunswick: International Society of the Learning Sciences.