Computer Supported Collaborative Learning Using CLARE: the Approach and Experimental Findings

Dadong Wan
Center for Information Technology & Management
Walter A. Haas School of Business
University of California
Berkeley, CA 94720, USA
Tel: (510) 642-4617
E-mail: dxw@haas.berkeley.edu

Philip M. Johnson
Department of Information and Computer Sciences
University of Hawaii
Honolulu, HI 96822, USA
Tel: (808) 956-3489
E-mail: johnson@hawaii.edu

ABSTRACT
Current collaborative learning systems focus on maximizing shared information. However, "meaningful learning" is not simply information sharing but, more importantly, knowledge construction. CLARE is a computer-supported learning environment that facilitates meaningful learning through collaborative knowledge construction. CLARE provides a semi-formal representation language called RESRA and an explicit process model called SECAI. Experimental evaluation through 300 hours of classroom usage indicates that CLARE does support meaningful learning, and that a major bottleneck to computer-mediated knowledge construction is summarization. Lessons learned through the design and evaluation of CLARE provide new insights into both collaborative learning systems and collaborative learning theories.

KEYWORDS: Computer supported collaborative learning; collaborative work; knowledge representation; knowledge construction; meaningful learning.

INTRODUCTION

Current computer-supported collaborative learning systems focus on improving shared access to information, people, and media. One example is virtual classroom systems, which range from plain e-mail, computer conferencing, and bulletin-board systems, to more specialized systems such as EIES [10]. Virtual classrooms allow learners to interact with their peers and instructors, as well as access online information in a manner independent of time and geographic location. They augment traditional classroom learning by removing the requirement for physical co-presence and by improving access to information and people.

A second example is hypermedia systems, such as Intermedia [25] and Mosaic [2]. These systems typically provide distributed mechanisms for structuring large information spaces. They also provide mechanisms for presentation and integration of various media formats, including text, graphics, voice, and video. Hypermedia systems improve access by removing the constraint of text-based interaction, broadening the scope of sharable information, and reducing the effort required to make information sharable.

Although virtual classrooms and hypermedia systems are successful in improving information access, they do not typically offer explicit mechanisms to help learners better interpret and assimilate the information, the context surrounding its creation and use, and the perspectives on it of the author or other learners. Simply improving information access without supporting learning leads directly to the problems of "information overload" and "lost-in-hyperspace". In the coming age of the Information Superhighway, it will be essential to explicitly support learning as well as access. Software systems must provide users with structural and process-level support on how to comprehend new information, how to relate new information to what they already know, and how to identify, compare, and integrate different interpretations of the same information. In other words, how to meaningfully learn in an environment of vastly improved information access.

This paper presents our approach to providing computational support for meaningful learning through a process of collaborative knowledge construction. The next section presents three major components of our approach: a representation language for meaningful learning called RESRA, a process model to guide the use of RESRA called SECAI, and a computational environment designed to support this language and process called CLARE. The following section describes our evaluation of this approach through about 300 hours of classroom usage. Analysis of the results indicates that the approach is effective in facilitating meaningful learning. It also reveals the strategies of knowledge construction used by learners, and several significant obstacles to meaningful learning within this paradigm. The following section briefly connects this research to related work in learning theory, cognitive psychology, and computer-supported cooperative work. The final section revisits the essential contributions of this research and outlines promising future research directions.

OVERVIEW OF CLARE

RESRA

RESRA, which stands for "REpresentational Schema of Research Artifacts," is a semi-structured knowledge representation language designed specifically to facilitate collaborative learning from scientific text, such as research papers. It has the following design goals:

- An organizational tool that allows incremental, fine-grained representation and integration of scientific artifacts;
- A mapping tool that highlights essential thematic features and relationships within and across scientific text, and that helps expose gaps and ambiguities in existing knowledge;
- A communication tool and a shared frame of reference that highlights similarities and differences between learners' points of view; and
- A tool for learning about the norms and conventions governing formal communication of scientific knowledge.

To achieve these goals, RESRA defines three types of conceptual constructs: *node primitives*, *link primitives*, and *canonical forms*. Node primitives represent discrete thematic features of the artifact, for example, *claim*, *concepts*, and *theory*. They also explicitly represent the learner's points of view in terms of *critiques*, *questions* and *suggestions*.

Link primitives describe relationships between thematic features represented by node primitives. For example, in a research paper, a *claim* is typically made with respect to a particular *problem* and must be supported by some *evidence*. In RESRA, these relationships are expressed as "*claim* $\stackrel{responds-to}{\longrightarrow}$ *problem*" and "*evidence* $\stackrel{supports}{\longrightarrow}$ *claim*," where *responds-to* and *supports* are link primitives.

The canonical form characterizes typical artifact-level thematic structures as a directed graph of RESRA node and link primitives. Research artifacts are classified into various types, such as *concept*, *empirical*, and *survey* papers. For each of these artifact types, a stereotypical RESRA structural graph can be identified. For example, in software engineering, one important type of research artifact is an "experience paper". Such artifacts report the experience of an organization with a software package or strategy, including the problem it attempted to solve or alleviate, the initial justification for adopting that software or strategy, and how the actual outcome compared to the expected result. In RESRA, a canonical form for experience reports could be expressed as shown in Figure 1.

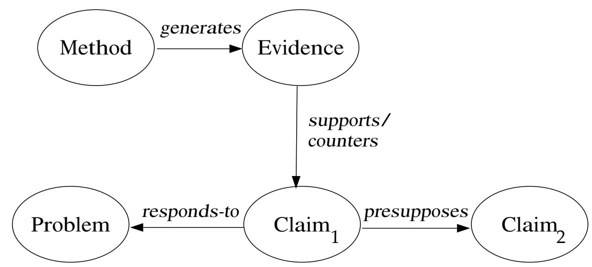

Figure 1: A RESRA canonical form for "experience paper"

Learners in software engineering may use the above structural model as a "template" to guide their interpretation and evaluation of all experience papers in that domain, and their efforts in constructing their own experience papers so that they may also conform to the same structure. In a group setting, such structural knowledge can be used as a shared framework for learners to engage in discussions about the content of related artifacts.

RESRA is based upon three premises: (1) Human knowledge can be represented in term of a small number of node and link primitives; (2) The use of these primitives to characterize scientific artifacts and subsequent group activities are a meaningful learning process; and (3) Different learners are likely to generate different representations of the same artifact; by comparing these representations, one can discern the similarities and differences in points of view held by individual learners.

Figure 2 provides a synopsis of all RESRA node primitives, while Figure 3 graphically depicts relationships between representative RESRA node and link primitives.

Node Type	Description	Example
Problem	A phenomenon, event, or process whose understanding requires further inquiry.	Meta-learning is not adequately supported by existing tools.
Claim	A position or proposition about a given problem situation.	Cleanroom engineering provides a viable solution in producing zero defect software.
Evidence	Data gathered for the purpose of supporting or refuting a given claim.	The use of cleanroom techniques led to a 5-fold reduction of defects in project Alpha.
Theory	A systemic formulation about a particular problem domain, derivable through deductive or inductive procedures.	Ausubel's theory of meaningful learning.
Method	Procedures or techniques used to generate evidence for a particular claim.	Delphi study; nominal grouping technique; waterfall software development model.
Concept	A primitive construct used in formulating theory, claim, or method.	Meta-learning; Knowledge representation.
Thing	A natural or man-made object that is under study.	Rock; Intermedia.
Source	An identifiable written artifact, either artifact itself or a reference to it.	An article by Ashton; the notes from Kyle's talk.
Critique	Critical remarks or comments about a particular claim, evidence, method, source, et al., or relationships between them.	Applications of cleanroom engineering appear limited to domains with well-defined requirements.
Question	Aspects of a claim, theory, concept, etc., about which the learner is still in doubt.	How does box-structured design differ from object-oriented design?
Suggestion	Ideas, recommendations, or feedback on how to improve an existing problem statement, claim, method, et al.	I would like to see cleanroom engineering used in some non-conventional domains, such as groupware.

Figure 2: A synopsis of RESRA node primitives.

Figure 4 shows an example use of RESRA that summarizes a seminal paper on software code inspection [8]. The *source* node in the upper left corner provides a reference to the artifact under study. The paper addresses one *problem*, against which three *claim*s are made. To support these *claims*, the author introduces three *concept*s and two *method*s. The latter are used to generate *evidence*, which in turn supports two out of the three *claim*s.

Several interesting observations can be made about Figure 4. First, the representation captures what is important in the research paper: it is not simply an outline of the paper but rather a map of its knowledge structure that reflects the learner's mental model about the author's intent. Second, as discussed later in this paper, different learners will derive quite different representations of this same artifact. Figure 4 is merely one of many possible representations. By comparing and contrasting these distinct representations, one can gain a better understanding of not only what this artifact is really about but perhaps more interestingly, how different learners interpret the same content. Third, also as described later, constructing these representations is non-trivial and requires significant learner effort, since RESRA node primitives, such as the *problem* node in Figure 4, may be only implied rather than explicitly stated by the author. Thus, the learner must infer these objects from the context of the paper, and relate them to other RESRA objects. Deriving such a representation is a meaningful learning experience because learners must ask themselves many deep-level questions: what is the claim(s) being made? With respect to what problem? Is a given theme a *claim* or *theory*? How are those themes related? Are there any "orphan" or unconnected themes? Answers to these and other related questions reveal what the artifact really means to individual learners.

SECAI

SECAI, which stands for "Summarization, Evaluation, Comparison, Argumentation, and Integration," defines an explicit process model for collaborative learning from scientific text. Figure 5 shows how these activities are related together to support collaborative knowledge construction. The world outside the concentric circles consists of various types of scientific artifacts, which constitute the raw material of learning. Metaphorically, collaborative learning with SECAI pulls learners from an external, isolated, and individual position inward toward an internal, integrated, and collaborative perspective on the artifact.

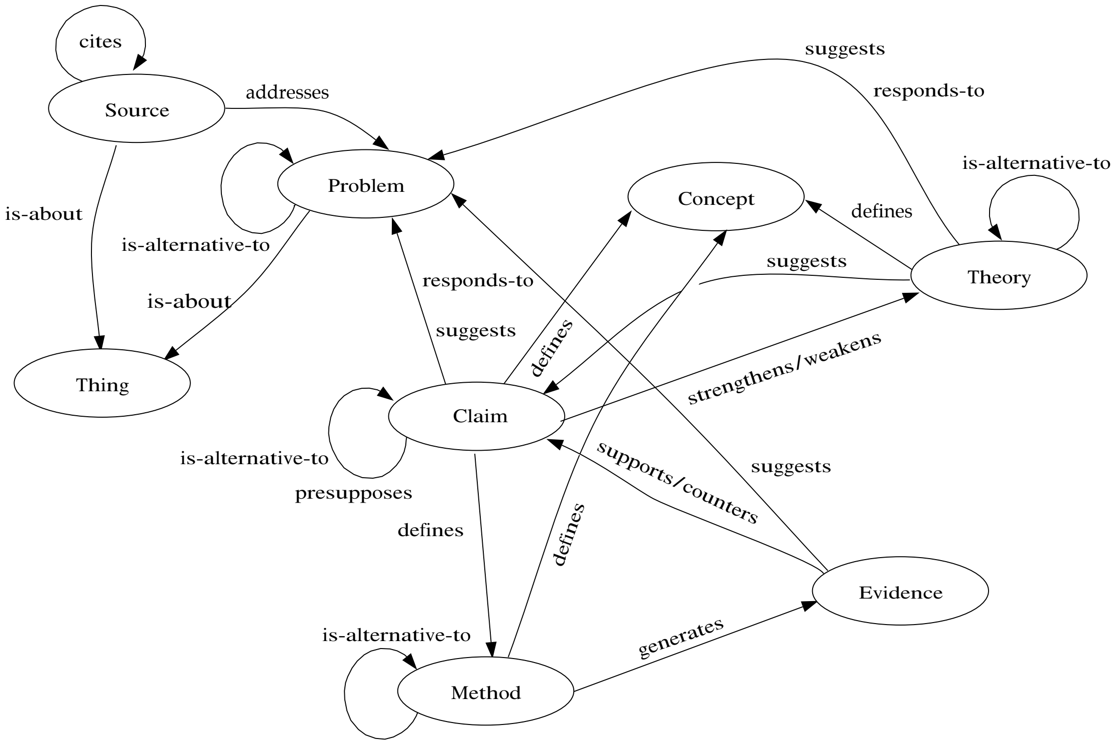

Figure 3: A graphical illustration of RESRA summarative node and link primitives and the relationships between them.

The first phase of SECAI is called *exploration*, which consists of two activities: *summarization* and *evaluation*. During this phase, learners derive a personal representation of the artifact and an evaluation of its content, both expressed in terms of RESRA. This phase is performed *privately* – learners are not allowed to see what other learners are doing or have done. This policy prevents learners from being distracted by each other's work or from free-riding off the work of others. The result of exploration is a set of representations similar to that shown in Figure 4 (except for the addition of evaluation nodes).

The second phase of SECAI is called *consolidation*, which consists of three activities: *comparison*, *argumentation*, and *integration*. During comparison, learners evaluate the similarities and differences between their representation and those of other learners. Comparison is done at three levels: the artifact level, the link primitive level, and the node primitive level. At the artifact level, learners compare their classifications of the artifact type, such as whether the artifact appears to be a *concept paper* or an *experience report*. In addition, they also compare their representations with respect to the canonical form selected for the artifact to see how their representations deviate from the standard one. At the link primitive level, learners compare their derived relationships, such as each learner's interpretation of the supporting evidence for a claim. Finally, at the node primitive level, learners compare their instantiations of nodes and the artifact content referenced by them.

Comparison activities provide a basis for *argumentation*. For example, suppose that John compares his representation to Jane's, and determines that his representation contains a *problem* node that is apparently missing from Jane's representation. He might then generate a *critique* node concerning Jane's representation, noting that it is missing an important problem raised in the artifact. Jane might respond by agreeing that her representation omitted an important problem. Alternatively, Jane might respond that one of her *problem* nodes in fact subsumes the problem noted by John. Another potential response might be to disagree with John's interpretation of the artifact content as a problem: that it was actually a method or claim of the research, as described in Jane's representation. This process of comparison and argumentation leads to an improved understanding of the meaning of the artifact. Perhaps as importantly, it reveals other learners' perspectives on the artifact.

The final step in the consolidation phase is *integration*, where learners create explicit links between their individual representations to improve their collective coherence and consistency. Going back to our hypothetical learners, if Jane realized that John had correctly identified a problem missing from her representation, she could integrate her representation by linking John's *problem* node into her representation in the appropriate places. Alternatively, if Jane believed that

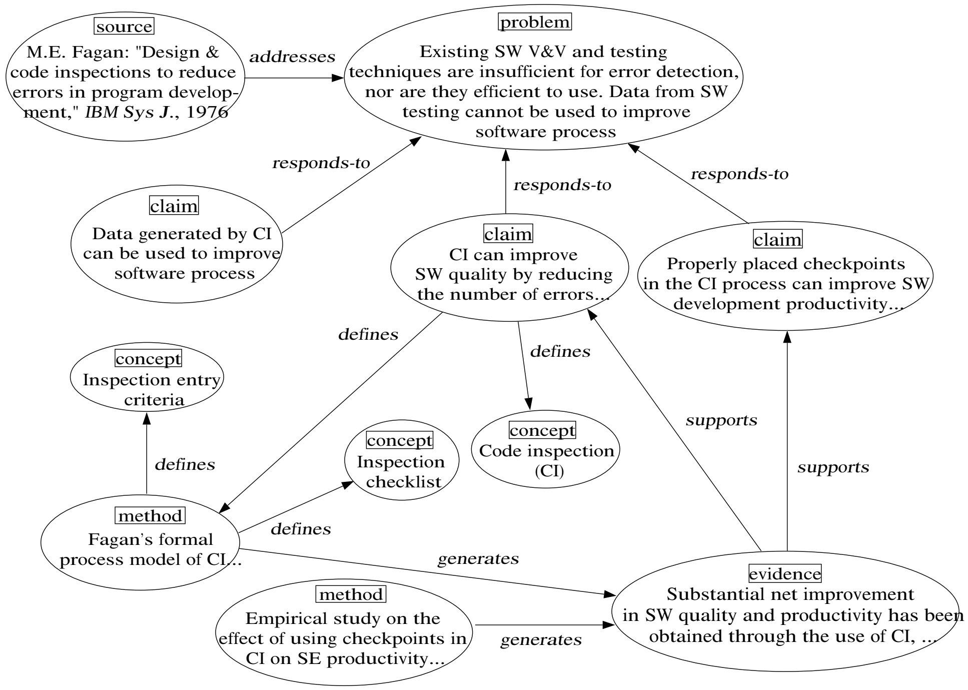

Figure 4: An expert's RESRA representation of Fagan's paper on code inspection.

one of her *problem* nodes subsumed the problem identified by John, she could integrate his representation by creating an *is-part-of* link between his and her nodes.

The large shaded arrows in Figure 5 indicate the direction in which SECAI "pushes" the group process: as learners proceed through the various activities, the level of collaboration among learners increases and, concurrently, a group knowledge base emerges. This dynamic knowledge base articulates both areas of consensus and areas of disagreements among a group of learners as they summarize, evaluate, compare, deliberate, and integrate their individual perspectives on the learning artifact.

CLARE

CLARE, which stands for "Collaborative Learning And Research Environment," is a distributed learning environment that supports SECAI and RESRA. It is a client-server system, running in a Unix/X-windows environment, and is built on top of the EGRET framework for exploratory collaboration [11, 12].

Use of CLARE begins by converting the scientific text to be studied into CLARE's internal hypertext format. Typically, the full text of the document is split up into smaller chunks called *source nodes*, each one corresponding to a physical section or subsection of the document. The current implementation of CLARE does not provide support for graphical images, so tables and figures are shown in the hypertext network as logical references and supplied to the learners as hardcopy documents.

Figure 6 shows a snapshot of the CLARE user interface during the exploration phase. The left window shows a source node corresponding to one section of the artifact under study. The node is connected to other source nodes via the *Up*, *Next*, and *Prev* links displayed on the first line of that node. Learners navigate through the scientific artifact under study by following these links.

To summarize a paragraph or any arbitrary block of text, the learner first highlights the text by dragging the mouse over it, and then selects the corresponding node type from the Summarize menu. CLARE creates a new node of the chosen type (e.g., *problem*, *evidence*) with default field template, and displays it in the lower right window. An explicit link is also automatically added between the selected text in the left window and the newly created node. The learner then provides annotative comments about the summarization in the **Description** field.

The above process is repeated until the learners believe they have fully summarized the document. (Summary nodes may also be created without reference to any text in the artifact.) Evaluation nodes are created analogously, although both source nodes and summary nodes may serve as targets of evaluation. The learner adds RESRA link primitives between two eligible nodes by choosing Link Mode from the Summarize menu. The upper right window in Figure 6 shows what the current learner has created so far for the

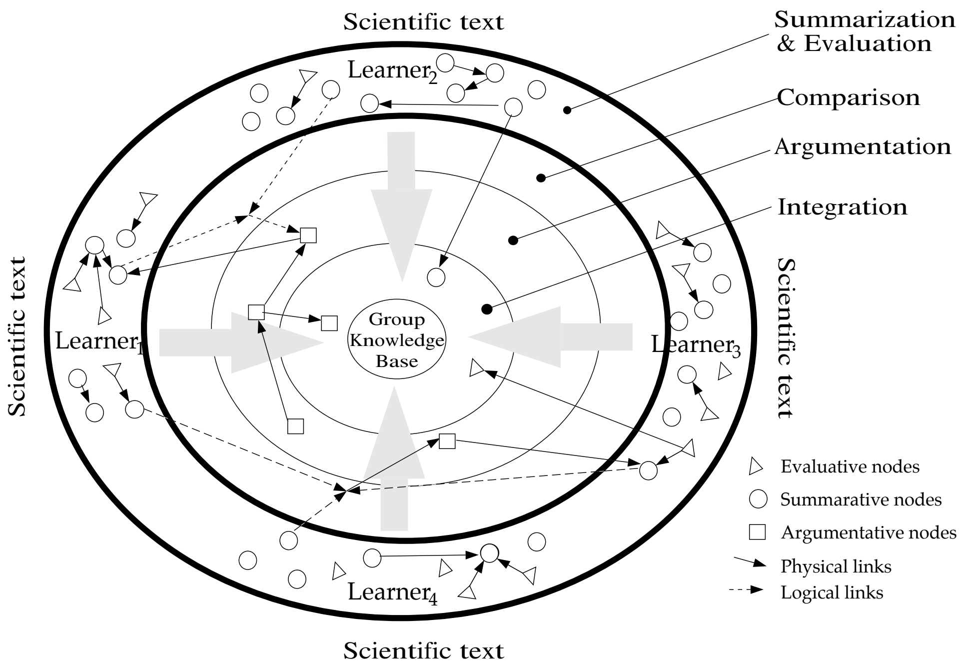

Figure 5: The SECAI process model for collaborative learning from scientific text.

current artifact.

Figure 7 shows a snapshot of the CLARE user interface during the consolidation phase. The upper left window displays a comparative view of the *problem* instances created by three learners (Peter, Cam, and Rose) during exploration. It highlights the differences and similarities between these learners with respect to their views on the problem the original authors attempt to address. Similar comparisons can also be made for the other RESRA primitives, for example, *claim*, *evidence*. To see the actual node instance corresponding to a given entry, the learner mouse clicks on the bold italicized text. If a learner then wants to know the precise place in the artifact from which the summarization node was derived, he can follow the link in the **Summarization** field.

To raise a question or make a critique on the *problem* node, one selects the corresponding RESRA node type from the Argument menu in the lower right window. This creates a new node instance of that type and links it to the *problem* node. When the creator of the *problem* node sees this new link, he can explain or defend his position using a similar procedure. He might also declare his *problem* as *similar-to* another node using the Integrate menu.

The upper-right window in Figure 7 summarizes the state of all learner's work. For example, Peter has created ten *claim* nodes, while Cam has seven, and Rose has six. Clicking on an entry generates a detailed listing for the corresponding user, from which individual node instances for any learner can be retrieved.

In addition to these computational services, CLARE also includes an instrumentation mechanism that unobtrusively gathers fine-grained process data. Each time the user performs a semantically interesting action, such as creating a new node or link, a timestamp representing this event and the time at which it occurred is recorded. The instrumentation also detects periods of idle-time for correcting elapsed-time calculations. This instrumentation provides vital data for answering such process-level questions as: what is the sequence in which learners visit nodes, and does this navigation strategy differ among learners? How much time do learners spend on each portion of the document? How much time do learners spend on each activity of SECAI? How are node creation and the link creation related procedurally? Answering these and other questions was the goal of a set of experiments we performed using CLARE, discussed next.

EXPERIMENTS AND FINDINGS

Experiments

Two sets of experiments were conducted to evaluate the effectiveness of CLARE as an environment for meaningful learning through collaborative knowledge construction. For this evaluation, three types of data were collected:

- *Assessment:* Gathered through a questionnaire administered after each experiment session;
- *Outcome:* Online CLARE database created during each

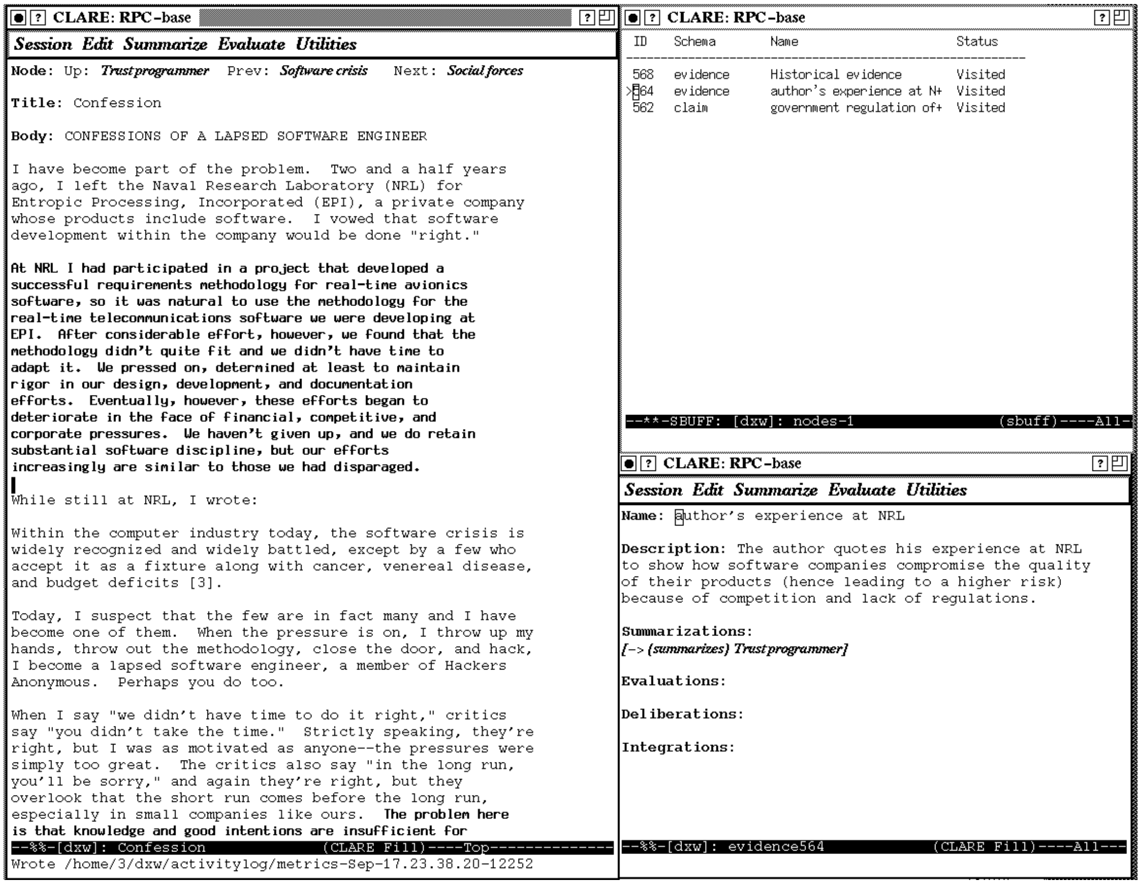

Figure 6: A user view of CLARE during the exploration phase. The left hand window contains a portion of the artifact under study. The lower right window contains an evidence node created by the learner. The upper right window summarizes what the learner has created during exploration thus far.

session; and

- *Process:* Gathered automatically through the built-in instrumentation mechanism.

The subjects were 24 computer science students who were enrolled in two software engineering courses in the Department of Information and Computer Science at the University of Hawaii in Fall, 1993. The task was collaborative analysis and deliberation of research papers in software engineering using CLARE. The first experiment involved 16 upper-level undergraduate students, who were evenly divided into four groups. The experiment was repeated three times with three different research papers. The second experiment involved 8 graduate students, who were evenly divided into two groups. This experiment was repeated twice with two different research papers. All experimental sessions lasted approximately one week. A few sessions lasted two to three days longer due to interruptions from other class activities.

The experiments were conducted between September and October, 1993. The subjects collectively accumulated about 300 hours of usage time, and created about 1,800 nodes with a total text size of nearly 400 kilobytes. A total of over 80,000 timestamps were gathered during these sessions. Figure 8 provides a short summary of the experimental data.

Exp.	Logins	Time (hrs)	Nodes	Size (Kb)
1a.	120	82.85	472	90.02
1b.	115	67.90	513	107.97
1c.	84	53.68	440	105.16
2a.	85	54.42	162	39.42
2b.	53	37.55	207	49.67
Total	457	296.40	1794	392.24

Figure 8: Selected CLARE statistics.

Results

Viability/usability of CLARE

The post-session survey responses from the subjects show that CLARE is a novel and useful collaborative learning tool: approximately 70% of learners indicated CLARE helped them understand the content of research papers in a way

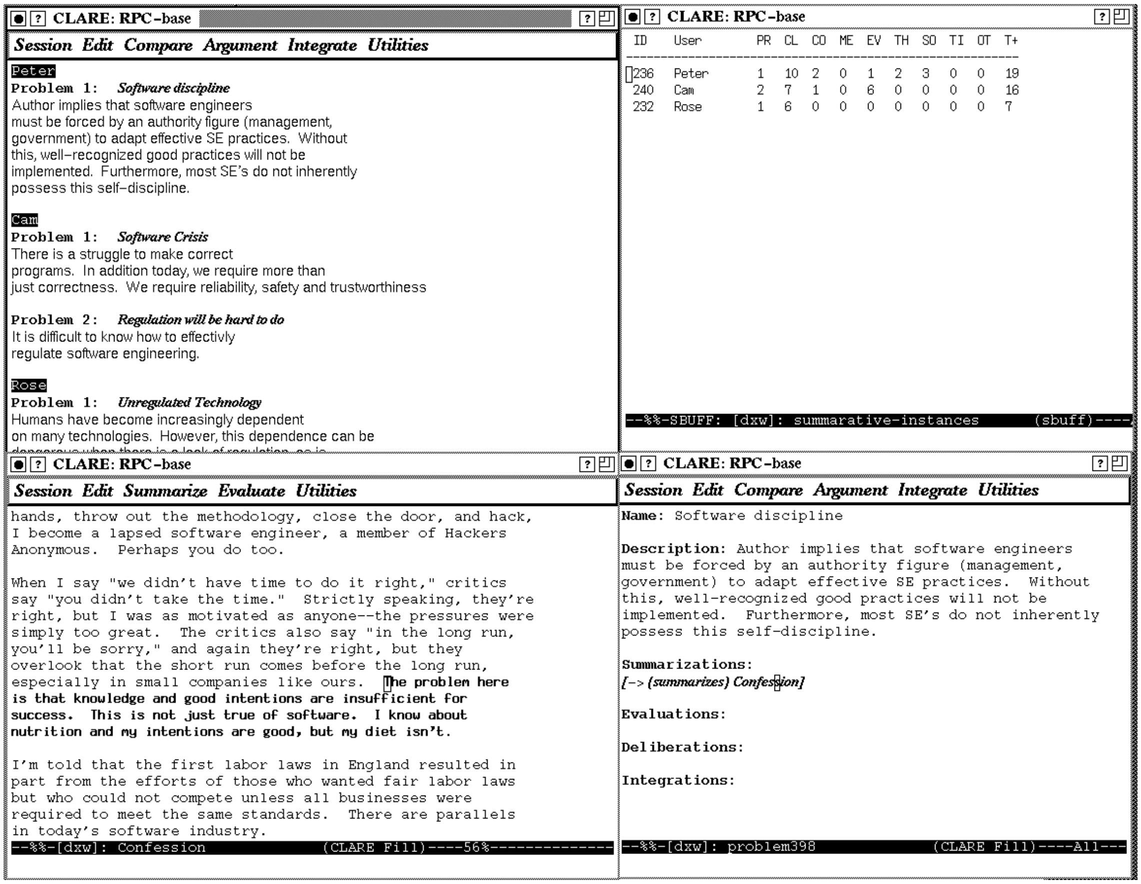

Figure 7: A user view of CLARE during the consolidation phase. The upper left window contains a comparative summary of the problems identified by each learner in the scientific artifact. One of the actual problem node instances is displayed in the lower right hand window. The lower left window displays the portion of the scientific text from which this problem was derived. The upper right window contains a summary of the activities of each learner.

not possible before, and nearly 80% of learners indicated that CLARE helped them understand their peers' perspectives in a way not possible before. Approximately 84% of the learners found that RESRA provides a useful means for characterizing the important content of research papers, and 90% of the learners agreed that RESRA helped expose different points of view on the same artifact.

Responses from the post-session survey also assessed the usefulness of individual components of our approach. The RESRA node primitives and the SECAI learning model were ranked the highest, assessed by 82% and 70% of the users, respectively, as "very" or "extremely" useful. The least useful features were the online examples, assessed by 25% of the users as "not" useful. RESRA canonical forms, the comparison mode, and link primitives were received mixed reactions from the user.

In addition to the empirical data, subjective responses to the approach were revealing. The following response from a subject shows that, in at least one instance, CLARE succeeded in fostering meaningful learning:

> *"... I don't quite know how to use it [CLARE] very well yet, but it really helped me get more out of the artifact we read. Without CLARE I would have just read the artifact and not really studied it or learned about the subject. CLARE made me look at the artifact from another point of view. That point of view was what is the author trying to tell me and how is the author trying to tell me that information ... Before I used CLARE I just read the artifacts. Now using CLARE I look for the meaning of the artifact and learn more about the subject..."*

Issues in CLARE-based collaborative learning

Detailed analysis of the outcome and process data revealed a number of interesting issues regarding collaborative learning using CLARE. These issues are discussed fully in [24], here we present four of the most significant: mis-interpretations of RESRA, failures in summarization, summarization strategies, and collaboration in CLARE.

First, RESRA was interpreted in a wide variety of ways among the learners. Despite the presence of hands-on training, detailed user documentation, and online examples, many subjects still seemed to fail to grasp the semantics of RESRA primitives, as evidenced by a substantial number of times in which RESRA nodes and links were used incorrectly. For example, though *theory* is defined as "a systematic formulation about a particular problem domain...", the following use of the primitive clearly does not satisfy this definition:

```
TYPE: theory
SUBJ: No single development improves
      the situation
DESC: No single development aids in
      improving the software problem,
      at least not with respect to
      productivity, reliability or
      simplicity.
```

Other typical errors in using RESRA include evidence nodes containing no evidence, suggestions containing no proposals from the learner, claims that are "neutral," evidence stated as claims, explanations or predictions identified as theories, problems treated as learner's disagreements with the author's claims instead of what the author attempts to address, and so forth.

Second, although subjects spent about 66% of their time on summarization, they frequently failed to adequately summarize the artifact by correctly identifying major themes and relationships and filtering out the minor ones. For example, Figure 4 provides an expert summary of [8] using 11 nodes. The 16 subjects analyzing this artifact generated an average of 19 nodes. Given this relatively large number, one would expect that all major themes would be covered, as well as a few minor ones. However, analysis of representations reveals that: (a) none of the 16 learners correctly identified the problem; (b) only seven learners correctly identified one or two of the three major claims; (c) only ten learners had the evidence right; and (d) only six learners had one of the two methods right. On the other hand, many minor themes of the paper were found in the learners' representations.

Third, the process data also reveals a set of stereotypical strategies used by learners in summarizing the content of an artifact. The strategies are characterized by the sequencing of summarative node and link instance creation:

- *Nodes only:* Create summarative nodes only. No attempt is made to connect them together using RESRA link primitives.
- *Nodes for an entire document, then links:* First create summarative nodes for the entire artifact, and then link them together.
- *Nodes, then links, but one section at a time:* Create summarative nodes for a single source node and/or its adjacent source nodes, and then create links between them; repeat the same process until all source node are summarized.
- *(2) first, then (3):* A combination of the first and second strategy. First create summarative nodes for the entire artifact, followed by a wave of link creation. Next, selectively create additional summarative nodes, immediately followed by the link creation.

Excluding the 36% of learners/sessions who adopted the first strategy, i.e., creating no summarative links, there was no noticeable correlation found between the strategy used and the quality of summarization.

Finally, in the SECAI model, explicit collaboration among learners takes place in the form of comparing their representations, deliberating reasonings behind them, and ultimately, integrating them into a coherent whole. Figure 9 shows an example collaborative representation network generated by four first-time CLARE users. Of a total of 92 nodes in the network, 34 were created during the argumentation phase. Most of these nodes (32 of 34) are *evaluative* in nature, which in turn can be categorized into two groups: pointing out the correct use of RESRA primitives and identifying ambiguities/inaccuracies in other learners' representations. In "critique642," for instance, Mary points out that, in "claim528," Scott has totally mis-interpreted the original authors' meaning. To assess the accuracy of Scott's representation, the process data shows that Mary in fact verifies the node content with the source from which Scott's node was derived.

Figure 9 also shows the presence of *constructive* (as opposed to *evaluative*) argumentation, in which learners do not merely critique or question each other's positions but engage in active knowledge-building by formulating new problems, proposing alternative claims, supplying additional evidence, and so on. In "evidence662," for example, Chris counters Mary's claim ("claim522") with new evidence.

Another noticeable feature about Figure 9 is the absence of integration activities, which turns out to be quite typical across CLARE sessions. A few learners elected to add explicit integrative links between their representations or vote for best representations. As a result, the group knowledge base consists of substantial amount of redundancy and inconsistency.

Discussion

The results provide a strong indication that RESRA, SECAI, and CLARE together provide a useful means for meaningful learning through collaborative knowledge construction.

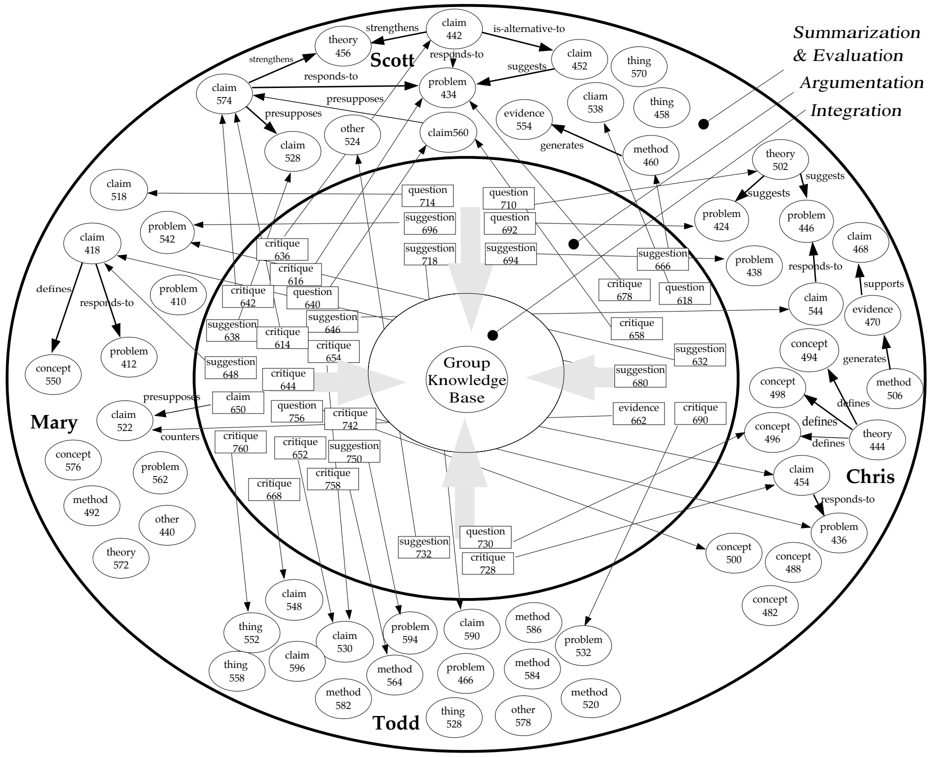

Figure 9: An example collaborative representation network by four CLARE users

Users indicated that the RESRA representation and the SECAI process model provided effective and novel support for representing the meaning of an artifact and comparing their interpretation to others. The results also indicated that users were not quite satisfied with the user interface of the CLARE prototype. Analysis of the experimental data provided new insight into the strategies used by subjects as they constructed their representations.

However, the results also yielded a surprising paradox: though users rated RESRA quite highly, they also made a substantial number of errors in its application. This raises the following question: is RESRA still useful for its intended purpose, despite individual variations in its interpretation and deviations from its intended usage?

The answer to this question appears to be a qualified yes, at least in the current context of use. It is possible, particularly at the node primitive level, for a learner to use a primitive incorrectly but usefully, simply because node and link primitives force the learner to analyze the artifact in terms of high level conceptual abstractions. Analyzing the artifact in this way provides a useful means to discover structural and content issues that might not be perceived otherwise. In addition, incorrect use of primitives also create collaborative opportunities, since the mistake might be identified by another learner. In many cases, the creation, identification, discussion, and correction of representation errors can facilitate meaningful learning both at the content level (by constructing an awareness of the true nature of the artifact under study) and at a meta-level (by constructing an awareness of the true nature of the representational language.)

However, incorrect use of RESRA node primitives can also raise a major barrier to effective use of the representation, since it can prevent effective use of RESRA link primitives. For example, an *evidence* either *supports* or *counters* a *claim*. At the node primitive level, it may not be of great consequence to misrepresent a *claim* as a *theory*. However, this mistake has significant consequences at the link primitive level, since the link primitives for claims will not be available. Thus, incorrect node primitive choices can inhibit the creation of links between nodes, a characteristic found in several subjects' representations.

During the design of RESRA, SECAI, and CLARE, we expected that the summarization activity would be rather straightforward and that most collaborative interaction would occur during argumentation and integration. The results surprised us: most of the subjects' effort and collaborative activity centered on summarization, and the summarization activity was not at all straightforward: subjects made substantial errors both in choice of RESRA primitives and in choice of the artifact content to summarize. We now realize that this seemingly straightforward process of reconstructing a scientific paper within a simple representation language poses an awesome challenge for many learners. The ultimate success of a collaborative learning environment has everything to do with how well it meets this initial challenge. We expect this problem will arise no matter what the nature of the representational language. Rather than hope for a "silver bullet" knowledge representation language, collaborative learning system designers should instead ensure that process-level mechanisms exist to overcome these breakdowns when they inevitably occur.

Finally, we want to point out the exploratory nature of the evaluation results discussed thus far: our primary goal was to assess the the viability and usability of the CLARE approach and to provide evidence and insights about what learners might do when confronted with such a novel learning environment. Through this study, we hope to provide a rich ground on which more rigorous experimentation and field studies on computer-supported collaborative learning can be formulated and performed.

RELATED WORK

CLARE is grounded in two theoretical tenets: social constructionism [4, 13] and the assimilation theory of cognitive learning [3, 17]. The former affirms the social nature of learning and the imperative of engaging learners in collaborative knowledge construction, as opposed to merely information sharing. It provides a philosophical foundation for the learning activities that CLARE supports. The latter is centered on the concept of *meaningful learning*, which defines learning as an ongoing process of relating new knowledge to what the learner already knows. Meaningful learning emphasizes explicit use of meta-knowledge to enhance human learning. Toward this end, two meta-cognitive strategies have been proposed: *concept mapping* and *Vee diagramming* [9, 17]. While the effectiveness of concept mapping is well supported empirically [6, 16, 18], its inadequacy as a collaborative learning tool and its lack of computerized support have directly prompted the current research.

RESRA is related to *schema theory* in cognitive psychology, which contends that human minds store and retrieve knowledge about the external world in terms of abstract chunks called schemas [21] and that the schema plays in an essential role in the selection, abstraction, interpretation, and integration of information [1]. RESRA is similar to some knowledge representation research in AI, particularly RA [22], which proposes an episodic representation for research literature.

A number of semi-structured representation schemes are found in the literature, for example, IBIS [14, 7], DRL[15], Toulmin's rhetorical model [23, 5]. RESRA differs from those schemes in that, among other things, it is fully integrated with an explicit process model (i.e., SECAI) that defines how the scheme is to be used.

CLARE falls into a special type of computer-based learning environments called *collaborative knowledge construction tools*, as contrasted with information sharing tools, such as EIES [10], Intermedia [25], and Mosaic [2]. CLARE is similar to CSILE (Computer-Supported Intentional Learning Environment) [20, 19] in that both systems reify a social constructionist paradigm and provide an environment conducive to collaborative knowledge construction. However, CLARE's representation scheme provides a meta-cognitive framework for collaborative learning, while CSILE's four thinking types ("I know," "high-level questions," "plan," and "problem") only allow learners to categorize their intentions. In addition, CLARE provides an explicit process model to control the application of the representation and the process of collaboration.

CONCLUSIONS AND FUTURE DIRECTIONS

This paper presents findings on a computer-based approach for supporting learning as a process of collaborative knowledge-building. The system, CLARE, differs from other learning systems in three important ways. First, it provides a semi-structured knowledge representation language that serves as a shared meta-cognitive framework to facilitate communication and collaboration among learners. Second, it defines an explicit process model of collaborative learning. Third, it implements fine-grained instrumentation mechanisms to gather detailed process data concerning the behavior of its users.

Analysis of experimental data confirms that CLARE is a novel environment that fosters meaningful learning. It shows that RESRA and SECAI provide useful structural and process-level guidance on how to collaboratively construct knowledge. In addition, analysis also reveals a number of issues for further research.

First, computer-supported collaborative learning is still a quite recent phenomenon for which no coherent theoretical frameworks yet exist. Many current learning theories, such as the ones on which CLARE is based, do not explicitly address such essential issues as how people develop shared mental models of the same artifact or problem situation, how to deal with differing terminologies for the same construct, and so on. As evidenced from the CLARE experimental result, these problems are intrinsic and also essential to collaborative knowledge-building, and cannot be answered by simply extrapolating proposals from individual-based learning theories. Hence, new and better theoretical explanations and guidance are called for.

Second, collaborative learning is not confined to classroom settings or scientific artifacts. Rather, it is part of every work situation in which artifact-based collaboration is required, such as software development, business proposal development, and so forth. Our long-term goal is to develop computer-supported environments that foster collaborative learning across task domains through tailorable structural and process-level support. RESRA and SECAI represent the first step in this direction.

Finally, we intend to apply CLARE's approach to collaborative learning as a means to assess the quality of the research/learning artifact and to help authors improve the quality of the artifacts they create. We hypothesize that a good scientific artifact contains a clearly-articulated knowledge structure and thus is easier to summarize using CLARE than a bad artifact. As a simple test of this hypothesis, we invite you—our readers—to contemplate a RESRA representation of this paper, and send us your assessments and discoveries.

Acknowledgments

The authors gratefully acknowledge the assistance of the other members of the Collaborative Software Development Laboratory in all phases of this research: Danu Tjahjono, Robert Brewer, Cam Moore, and Rosemary Andrada. Support for this research was partially provided by the National Science Foundation Research Initiation Award CCR-9110861.

REFERENCES

1. J.W. Alba and L. Hasher. Is memory schematic? *Psychological Bulletin*, 93:203–31, 1983.
2. Marc Andreessen. NCSA Mosaic technical summary. Available via anonymous ftp from "ftp.ncsa.uiuc.edu:/Web/mosaic-papers/mosaic.ps.Z".
3. David P. Ausubel. *The psychology of meaningful verbal learning*. Grune & Stratton, 1963.
4. Peter Berger and Thomas Luckman. *The Social Construction of Reality: a Treatise in the Sociology of Knowledge*. Doubleday, 1966.
5. Violetta Cavalli-Sforza, Gareth Gabrys, Alan M.Lesgold, and Arlene Weiner. Engaging students in scientific activity and scientific controversy. In Kishore Swaminathan, editor, *Proceedings of 1992 AAAI Workshop on Communicating Scientific and Technical Knowledge*, pages 99–106, San Jose, 1992.
6. Joseph W. Cliburn, Jr. Concept maps to promote meaningful learning. *Journal of College Science Teaching*, pages 212–217, February 1990.
7. J. Conklin and L. Begeman. gIBIS: A hypertext tool for exploratory policy discussion. *ACM Transactions on Office Information Systems*, 6(4):303–331, October 1988.
8. Michael E. Fagan. Design and code inspections to reduce errors in program development. *IBM System Journal*, 15(3):182–211, 1976.
9. D.B. Gowin. *Educating*. Cornell University Press, 1981.
10. S.P. Hiltz. Collaborative learning in a virtual classroom: Highlights of findings. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work*, pages 282–290, 1988.
11. Philip M. Johnson. Supporting exploratory CSCW with the EGRET framework. In *Proceedings of the 1992 Conference on Computer Supported Cooperative Work*, November 1992.
12. Philip M. Johnson. Experiences with EGRET: An exploratory group work environment. *Collaborative Computing*, 1(1), 1994.
13. K.D. Knorr-Centina. *The manufacture of knowledge: an essay on the constructivist and contextural nature of science*. Pergamon Press, Oxford, 1981.
14. W. Kunz and H. Rittel. Issues as elements of information systems. Technical report, Institute of Urban and Regional Development, University of California at Berkeley, 1977.
15. Jintae Lee and Kum-Yew Lai. What's in design rationale? *Human-Computer Interaction*, 6(3,4):251–280, 1991.
16. J.D. Novak. Concept mapping: a useful tool for science education. *Journal of Research in Science Teaching*, 27(10):937–49, dec 1990.
17. Joseph D. Novak and D. Bob Gowin. *Learning how to learn*. Cambridge University Press, 1984.
18. Wolff-Michael Roth and Anita Roychoudhury. The social construction of scientific concepts or the concept map as conscription device and tool for social thinking in high school science. *Science Education*, 76(5):531–557, 1992.
19. Marlene Scardamadia, C. Bereiter, C. Brett, P.J. Burtis, C. Calhoun, and N.S. Lea. Educational applications of a networked communal database. *Interactive Learning Environments*, 2(1):45–71, 1992.
20. Marlene Scardamalia and Carl Bereiter. Technologies for knowledge-building discourse. *Communications of the ACM*, 36(5):37–41, May 1993.
21. Neil Stillings, Mark Feinstein, Jay Garfield, E.L. Rissland, D.A. Rosenbaum, S.E. Weisler, and L.Baker-Ward. *Cognitive science: an introduction*. MIT Press, 1987.
22. Kishore Swaminathan. *RA: a memory organization to model the evolution of scientific knowledge*. PhD thesis, University of Massachusetts at Amherst, 1990. also COINS Technical Report 90-80.
23. S. Toulmin. *The use of argument*. Cambridge University Press, 1958.
24. Dadong Wan. *CLARE: A Computer-Supported Collaborative Learning Environment Based on the Thematic Structure of Scientific Text*. PhD thesis, University of Hawaii, Department of Information and Computer Sciences, 1994.
25. Nicole Yankelovich, Bernard J. Haan, Norman K. Y. Meyrowitz, and Steven M. Drucker. Intermedia: the concept and the construction of a seamless information environment. *Computer*, January 1988.