Towards an On-Line Student Assistant within a CSCL Framework

Fernando Ramos-Quintana, M. Dolores Vargas-Cerdán, J. Rafael Rojano-Cáceres
Tecnológico de Monterrey, Campus Cuernavaca Autopista del Sol, Xochitepec, México
Email: fernando.ramos@itesm.mx, dvargas@uv.mx, rojano@uv.mx

Received July 20th, 2012; revised August 18th, 2012; accepted August 30th, 2012

The CSCL framework of this work aims at reinforcing knowledge by rebuilding and representing it synthetically in form of a network of concepts, which is built through exchanged messages using a chat and a graphical tool. We have found that during the collaborative learning sessions important troubles bring about deadlocks affecting the learning process. The hypothesis of this work states that significant troubles are expressed by students in form of assistance requests originating behavior patterns associated with certain phases defined within the collaborative learning sessions. Analysis of dialogs of real collaborative learning sessions confirmed this hypothesis, whose results point towards the development of an on-line assistance system that aims at breaking the deadlocks caused by troubles expressed in the form of assistance requests. Derived of an analysis of 55 dialogs of a course of databases, we defined three phases characterized by particular assistance requests occurring during the collaborative learning sessions: a) a bootstrap phase characterized by requests about doubts of the use of tools; b) a work phase, wherein requests are related with the construction of the network concepts; c) the goodbye phase wherein requests reflect doubts about the final deliverable network.

Keywords: Computer-Supported Collaborative Learning; Dialogues; Behavior Patterns

Introduction

This work is situated within a CSCL environment that uses dialogs to build knowledge synthetized and represented by a network of concepts (NoC). This CSCL environment aims at encouraging collaborative learning attitudes and the development of cognitive skills to retain and build knowledge. The knowledge about the topic under study is synthetized in the form of a NoC represented by a graph similar to Petri Nets.

The communication between students to collaborate in the construction of the network of concepts is established through a dialog, which can be formally defined as follows:

\[ \text{Dialog} = m_1 < m_2 < \cdots < m_{n-1} < m_n \]

where, the symbol \(<\) represents the temporal term “before”; \(m_i\) represents any message, and \(i\) and \(n\) are integer numbers. Therefore, the dialog can be seen as a totally ordered sequence of exchanged messages whose goal is to build, in a collaborative way, the knowledge of a topic expressed synthetically by a NoC.

The exchanged messages contain not only drawing actions that aim at linking concepts to form the network, but also doubts about the use of the tool, the way of linking the concepts and general questions about the final product to be delivered. The doubts about the use of the tool and the way of linking concepts represent the most important causes of deadlocks that affect the performance of the collaborative learning process. We have observed that as the dialog evolves, some important behaviors associated with requests of assistance are related with the causes of deadlocks mentioned before.

In this article, we are focused on the analysis of messages of dialogs to determine a set of behavior patterns that can aid to identify the required assistance. These results are being used to develop an on-line assistant system, currently under development, that can provide students with aids to break the deadlocks. This assistant system is out of the scope of this article.

We have discovered patterns that reveal behaviors by analyzing 55 dialogs of 55 pairs of students. These 55 dialogs, which constitute a total of 4910 exchanged messages, are derived from three different semesters, two groups by semester, of a database course of an undergraduate program in Computer Science, based on this analysis we have defined different types of assistance requests which cause deadlocks and consequently damage the performance of the collaborative learning process. Thereby, the types of assistance requests were defined within three phases related with behaviors of students occurring during the construction of the network: a) the initial phase is characterized by assistance requests mainly related with the use of the tool; b) in the intermediate or work phase, the students requested mostly assistance about the construction of the NoC; c) the final phase is characterized by requests about details of the built network and the product to be delivered.

The remaining sections of this paper are organized as follows: the second section reviews relevant work related with our work in the domain of CSCL systems; the third section describes the CSCL system within which this investigation is carried-out. The construction of network of concepts is specially treated in this section; the fourth section presents an analysis and discussion of qualitative and quantitative aspects of the messages, which were used for the analysis that aimed at determining behavior patterns that affect the collaborative learning processes within the context of a course of databases; finally, the fifth section shows the conclusions.
Relevant Related Work

CSCL Systems

Dillenburg and other people coincide in defining collaborative learning as a situation within which two or more people learn or attempt to learn something together (Dillenburg et al., 1996; Dillenburg, 1999; Collazos, 2001). It improves the individual learning and provides the students with the opportunity to develop cognitive skills while they develop social abilities (Barnes, 1977). Roschelle (Roschelle et al., 1995) affirms that collaboration is “the mutual engagement of participants in a coordinated effort to solve a problem together”.

The benefit of the collaborative approach for learning lies in the processes of articulation, conflict and co-construction of ideas occurring when working closely with a peer. Participants in problem-solving situations have to make their ideas explicit to other collaborators, disagreements, justifications and negotiations, helping students to coverage to a common object of shared understanding (Crook, 1996).

One of the most important joint projects of computer science and education research is determining how to devise and deploy software tools to support classroom instruction. One approach is related with computer supported collaborative learning (CSCL) systems, which advocate group collaboration to promote learning among the students.

However, the benefits of collaborative learning are only achieved by active teams that function well (Soller, 2001). In addition, adequate technology tools and methodologies should be constantly improved to enrich the collaborative learning environments: in a review by Soller (Soller et al., 2005) they studied tools and methodologies that support collaborative learning interactions; argumentation in collaborative learning systems can help to facilitate the communication between participants as argued in a review of computer supported argumentation made by Scheuer et al. (Scheuer, 2010); an essay of the relationship between technology and collaborative learning performance is carried-out in Stahl (Stahl, 2006); such improvements can be related with the development of network architectures (Bote, 2004).

The analysis of collaborative interactions aims at monitoring student’s activity and assessing the performance of learning: Beatriz Barros and M. Felisa Verdejo (Barros and Verdejo, 2000) analyzed student’s interactions to improve collaborations; Juan (Juan 2008) developed an information system to monitor the student’s activity in online collaborative learning; Putambekar (Putambekar, 2006) aimed at knowing whether as the collaborative interactions evolve new ideas raise or the built knowledge is enriched; in a study made by Yang and Chang (Yang & Chang, 2012) it has been found that interactive blogs are associated with positive attitudes towards academic achievement.

Assistance systems to help students in case of stuck situations have been one of the most important concerns in CSCL. The work presented in this paper belongs to these kinds of systems. Dialogs generated in collaborative learning systems have attracted an important number of researchers, who have been interested in their analysis as source of data to gain insights into the quality and performance of the collaboration process. Other researchers like Pilkington (Pilkington, 2001) consider that the analysis of dialogs can help bridge the gap between the empirical evaluation of interaction and the design of Intelligent Educational Systems (IESs) capable of interacting with their users; Woojin Paik et al. (Paik, 2004) aimed at developing intervention techniques to identify and remove obstacles of online learning groups. The final goal is to build an automatic system able to monitor the activities of the online learning group members to alert the instructors when the members encounter barriers; In web-enhanced courses, instructors would like to review the discussions of students to understand the kind of contributions, monitor their progress or verify if they need assistance or guidance (Painter et al., 2003).

Some works consider the intervention of agents to support the collaborators by providing aids in stuck situations or improving the participation of students. In (Kim, 2007) speech act analysis was carried-out to assess the effect of instructor intervention on student participation, resulting in an increase of the number of exchanged messages made by the students. However, one of the most complex problems to assist the participants in on-line mode is to be able of detecting the different categories of messages, and thus defining the kind of assistance that could correspond to that one requested by a particular message. The following approaches, most of them based on the analysis of speech acts derived from dialogs under study, could potentially be adapted and applied to contribute to solve several aspects of this problem: the Cohen’s work (Cohen, 2004) aimed at detecting many categories of messages with high precision and moderate recall by using text-classification learning methods; the Feng’s work (Feng, 2006) is centered on detecting which messages in a thread contains the focus of the conversation, localizing what messages are related with the subject, recovering past conversations and using them to solve doubts. This work integrates studies of conversational speech acts, an analysis of message values based on poster trustworthiness and an analysis of lexical similarity; in (Ravi & Kim, 2007), speech act classifiers were developed which were able to identify whether a message contains questions or answers; the purpose of a subsequent Kim’s work (Kim, 2008) was to develop tools that could automatically assess student participation and promote interactions by sending responses to student messages. These software tools apply data mining and information retrieval techniques for guiding student discussions. Thus, past student discussions of related courses are used and the retrieved information is presented on the discussion board; in (Kim et al., 2005) was developed a course ontology that represents generic components of distance education courses and a query ontology that describes types of student’s queries and requests. The system can map student queries to relevant course materials and the results are sent to the students. In case of non-appropriate mappings or the sent materials do not satisfy the student requests, the case is sent to the instructor’s attention; in (Caballé, 2008) is proposed a model based on data analysis from online collaborative interactions towards an automatic assessment in real time. The use of machine learning approaches is suggested to label automatically the messages; in an interesting approach by Geo (Geo et al., 2011) it was examined whether interaction patterns could be classified automatically by using a state transition model to identify successful versus unsuccessful student Q & A discussions, and classification of threads as successful/unsuccessful using the state information.

The analysis of messages to determine what assistance has been requested could be derived from only one message or from a sequence of them. Nevertheless, to precisely the requested assistance based only on one message, and moreover only considering key words belonging to it, could be result erroneous because a correct interpretation of a request needs contextual
information, which is usually obtained by considering a sequence of messages but not only one. This approach has associated an important complexity, because one isolated message could have several interpretations, thus two or more messages together result in the product of individual interpretations of the messages being considered. For instance, if one message is associated with three illocutive acts, in case of an analysis based on speech acts, then this message would have three potential interpretations. Such problem could require dealing with more than one message to facilitate the interpretation. In order to simplify the exercise, let’s suppose now that three messages are being considered to give an adequate contextual interpretation; and let’s also suppose that each one of the messages has three possible interpretations. Thus, the product of these three messages results in at least nine possible interpretations and thus the interpretation ambiguity starts to be a real problem. Moreover, this exercise supposes that the participants write the messages with a correct syntax and semantic way, which is not commonly the case, at least in our Mexican reality. The modern technologies of communications, such as the ones using in collaborative learning systems supported by computers, have considerable modified the way of writing in such a way that is very hard to associate a formal linguistic structure with dialogs established by students during a collaborative learning process. Therefore, the interpretation ambiguity increases significantly and then the application of a formal approach, such as the analysis based on speech acts, aiming at interpreting the intention of messages becomes a very complex task to be accomplished. In addition, as exposed by Kim (Kim, 2008), the discussion data from students, especially in undergraduate programs, is highly incoherent and noisy. The raw data includes humorous messages and personal announcement as well as technical questions and answers.

Owing to the problems exposed before, we have adopted another approach, different than those based on the interpretation of speech acts, towards the on-line intervention in the collaborative process when the students request assistance in stuck situations. The study presented in this paper deals with one relevant step of the whole objective, which is to assist student requests in on-line mode to unstuck deadlock situations. This step aims at characterizing the kind of assistance requests associated with different phases of the learning process during a work session. We have proven, in a previous work (Ramos-Quintana et al., 2008), the existence of a pattern composed of three phases during a collaborative work session: a bootstrap phase, a work phase and a goodbye phase. This issue is discussed later in the forth section.

The Computer Supported Collaborative Learning (CSCL) System

The CSCL model is supported by a collaborative computer tool called Collaborative Distributed Tool (CDT) (Rojano-Caceres et al., 2010) developed by the research group on Computer Supported Collaborative Learning from the Tecnológico de Monterrey Campus Cuernavaca and the Universidad Veracruzana. This tool, which can be accessed at http://collaborativelearningframework.net/, is integrated by three elements: 1) A chat tool that serves to build the dialog whose objective is to generate the drawing actions to build the NoC; 2) A graphical working space, where the network of concepts is built, and; 3) A bottom-up methodology to build the NoC.

Figure 1 illustrates the interface of CDT. As we can see, two
students communicate through the chat to build the dialog. A shared area allows students to build the NoC by drawing geometric figures denoting nodes and directed arcs (represented by arrows). Two kinds of nodes are available: nodes in form of circles represent concepts to be linked; nodes in form of rectangles represent relations between concepts. Directed arcs link circle-nodes with rectangle-nodes and vice-versa. Below we explain how a NoC is built, as well as the role of nodes and directed arcs in this construction.

The Network of Concepts: One of the main abilities when the retention or acquisition of knowledge process is taking place is to synthesize and represent such knowledge in a structured way by linking the concepts involved in a correct semantic way. For this purpose, the development of the skill of analysis will allow to extract the most relevant underlying concepts and establish relations between them in order to get a structural representation of knowledge being learned. This structural representation is called Network of Concepts (NoC) in this work. This way of working leads students to build knowledge from the isolated underlying concepts to structural concepts by developing at the same time different abilities such as analysis, synthesis, construction of structures, abstraction and generalization, which are very important in the process of solving real problems.

The NoC starts as follows to be built: underlying concepts are represented by circle-nodes and rectangle-nodes serve to link two circle-nodes through directed arcs represented by arrows. Two kinds of rectangle-nodes are available: black rectangle-nodes represent dynamic links and white rectangle-nodes represent static links. We mean by dynamic links those that imply actions of ‘execution’, ‘processing’ or ‘transformation’; in the case of transformation a concept \( c_1 \) is transformed into a concept \( c_2 \) through a dynamic relationship \( r \), as expressed by the example of the rule (1) shown below.

\[
C_1 \xrightarrow{r} \text{metamorphosis} C_2 = \text{frog}
\]  

Semantically, this rule means: if a tadpole undergoes a metamorphosis, then it will be transformed into a frog.

Meanwhile, static relationships denote some reference or characteristic to another concept that does not imply a transformation. For example, it can be used to stay the fact that ‘the sun is a star’, such representation is straightforward, which is built by linking the concept of “Sun” with the concept of “Star” by the relationship “is a”, in Figure 2 we present graphically this example.

The rest of the network is built following the same procedure. For example, in the theme of Relational Model from the Database subject we can get important concepts such as relations, tuple or rows and grouping, those can be represented as shown in Figure 3.

The Methodology for the Construction of the NoC

We exemplify this methodology for the reinforcement moment, wherein a topic has been learned previously and we aim at improving the retention of isolated underlying concepts and
how they are structurally linked. At the beginning of the process, the teacher provides the student with a topic to be reinforced, after carrying out an important effort of analysis the students try to list the underlying concepts of the topic (the first level). Thus, at this level, the development of the analysis skill is prioritized; at the next level, underlying concepts are linking by pairs through dynamic or static rectangle-nodes. The skills to be developed at this level are, once again the analysis, then synthesis and the construction of structures composed of two linked underlying concepts; finally, they link the pairs until getting the whole NoC. At the last level, the skills to be developed are: the analysis, the synthesis, the construction of structures, the abstraction and the generalization.

Qualitative and Quantitative Analysis: A Case Study

Description: This case of study took place at the Statistics and Computer Science Faculty of the Universidad Veracruzana in Mexico. This investigation has been made by analyzing the contents of dialogs of 55 pairs of students in the field of a database course in the undergraduate level. The 55 dialogs are derived from several courses of three different semesters. In a previous work (Ramos-Quintana et al., 2008), we have reported a general pattern composed of three sequentially ordered phases: 1) the handshake phase (now renamed “the bootstrap phase”), because the students exchanged messages not only to say hello, but also messages that denote a kind of setup to start working in the construction of the NoC; 2) the work phase, where the construction of the network takes properly place; 3) the closing of the dialog, where the students say good bye.

In this work, we have characterized those messages belonging to each phase that reveal requests, implicit and explicit, of assistance. The following messages are examples of implicit or explicit requests of assistance:

- We mean by explicit messages of assistance requests those messages containing a question that reveals an explicit trouble that entails deadlocks in the learning process.
- We mean by implicit messages of assistance requests those messages containing an explicit trouble that entails deadlocks in the learning process.

Type of explicit messages: 1) How do I do to write inside the circles? 2) But, how do I put arrows? 3) What is the purpose of the arrows?

Meanwhile, in the implicit messages, there are not questions explicitly expressed, but the messages express also implicitly troubles that deadlock the learning process.

Type of implicit messages: 4) I’m trying to find out how to link nodes!!! 5) I guess that the direction of the arrows of the items inside the tables should point to the opposite direction.

The purpose of this work is not to represent automatically the meaning of messages. Instead, we make a not automatic analysis to extract those messages containing implicit or explicit request of assistance to characterize and associate them with the different phases of a learning session described before.

Type of Messages of the Bootstrap Phase

During the bootstrap phase, the exchanged messages mainly concern some salutations, questions to know whether the couples have enough information to start and questions, highlighting doubts about the use of the computer tool and the objective of the task to be carried-out. Some examples of messages representing the bootstrap phase are the following: Hi, do we insert the list of concepts??? I am ready; Do you have some concepts??? What’s up; What concepts you got; Do you know how to insert arrows to link two nodes?? What do we have to do?? Ok, let’s start to work; why should we work with the most important concepts?? I’m trying to test the communication…; Tell me how many concepts you have, I got five.

Type of Messages of the Work Phase

During the work phase the troubles are related with the construction of correct structures, for instance, why a node X should be linked with a node Y and not with the node Z, the sense of the arrow, etc. Some examples of typical messages associated with the work phase are the following: Could we put the transition?? But, we have to build the whole network?? We have to get more concepts; Is it ready now??? Do you have any another idea???? Do you know which is the sense of this arrow?? Let me see if I remember something else; The tuples form a database?? Do you know why the node X should be linked with the node Y…??

Type of Messages of the Goodbye Phase

In the goodbye phase the students have doubts about whether the NoC is already finished or about the destination of the final deliverable product, etc. Some examples of messages associated with the goodbye phase are the following: Here it go…!!! I guess the network is ready; I’m not completely convinced… But… is it ok like that??? Can we already leave; So, what else; Do you think it is ready to be sent…??? What’s the e-mail address to be sent?? But, what you think about this? Let’s make the last touch.

Quantitative Aspects

The next analysis illustrates how much time has the couples taken for each phase. The analysis has been done by considering 55 dialogs, for a total of 4910 exchanged messages. Each dialog corresponds to a couple. For instance, Table 1 below shows ranges of time consumed by students during the bootstrap phase. Thus, most of the students, 45 couples from 55 (81.8% of couples), fell in the range between 30 seconds to 6 minutes in the bootstrap phase. Meanwhile, 6/55 (10.9%) fell in the range between 7 to 13 minutes, and finally 4/55 (7.2%) in the range between 14 to 22 minutes.

Table 2 shows ranges of time consumed by students during the work phase. We can see that most of the couples 31/55

<table>
<thead>
<tr>
<th>Table 1. Bootstrap phase.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranges of time in minutes</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>From 0.5 to 6</td>
</tr>
<tr>
<td>From 7 to 13</td>
</tr>
<tr>
<td>From 14 to 22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Work phase.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranges of time in minutes</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>From 15 to 28</td>
</tr>
<tr>
<td>From 29 to 46</td>
</tr>
<tr>
<td>From 47 to 60</td>
</tr>
</tbody>
</table>
(56%) during the work phase fell in the range between 29 to 46 minutes to complete this phase. Meanwhile, 14/55 (25%) fell in the range between 15 to 28 minutes. And 10/55 (18%) couples took between 47 to 60 minutes.

Table 3 shows ranges of time consumed by students during the goodbye phase. We can verify that most of the couples 44/55 (80%) fell in the range from 30 seconds to 6.5 minutes to complete this phase. Meanwhile, 6/55 (10.9%) took between 6.31 to 10.3 minutes and 5/55 (9%) fell in the range between 10:31 to 14:31 minutes.

Qualitative Aspects

As mentioned before, a total of 55 dialogs were analyzed, which totalized 4910 messages. 1558 messages from a total of 4910 were messages that requested assistance, which represents 31% of these messages, this amount is significantly important because it represents 1/3 of the whole set of messages. The total of messages denoting a request of assistance for the three phases is qualitatively shared as follows: 138 (bootstrap phase), 1240 (work phase), 180 (goodbye phase). Thus, from this 31%, 25% belongs to the work phase (1240/4910), that is, 80.6% of the total of messages that reveals assistance request. Due to the fact expose before, we can conclude that most of the request requiring a help is concentrated in the work phase, that is, in doubts about the semantic construction of the network of concepts, as expected.

Table 3.

<table>
<thead>
<tr>
<th>Ranges of time in minutes</th>
<th>Number of couples</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>From 0.3 to 6.5</td>
<td>44</td>
<td>80</td>
</tr>
<tr>
<td>From 6.6 to 10</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>From 10 to 14</td>
<td>5</td>
<td>9</td>
</tr>
</tbody>
</table>

Figure 4 shows an average curve denoting behavior patterns for each phase. We can observe a certain increase of the curve at the beginning of the bootstrap phase, which is characterized by messages of assistance requests related with the use of the tool, the preparation of materials to start with the work phase and some expressions denoting salutations. At the end of this bootstrap phase these kinds of messages tend to disappear. However, the curve tends to increase right after the point where the work phase begins. The work phase has its highest peak within the period of time between 32 to 35 min. In the neighborhood of this period of time the students suffered most of the troubles related with the semantic construction of the network. After the neighborhood of this interval of time (32 to 35), the curve tends at decreasing until the end of the phase. At this point, the goodbye phase starts to increase but the peak is not enough relevant to be considered as a serious cause to provoke a real deadlock. The kinds of messages in this phase are related with doubts about the finished work and also doubts about the destination of the deliverable work, among the most important.

Conclusion

Some efforts have been carried out to provide CSCL systems with automatic capacity of assistance or supervision of student’s performance during the learning process interaction. We are currently working towards building such computer assistant system. However, the development of this system needs some important steps to be achieved. One of them aims at determining the adequate moment to assist the students along with the kind of assistance they need: this work showed the results concerning the study of both problems.

During the collaborative learning process an important number of troubles can bring about deadlocks, due to several reasons, which put at risk the performance of the learning process. Working under the collaborative learning framework described in this article, we have confirmed that students request repeti-
tively for assistance to solve some troubles, varying the fre-
quency of such requests in depending on the phases, described before, of the learning process. In this work, we aimed at de-
termining quantitatively the number of requests for assistance 
and discovering qualitatively the type of request of assistance, 
which can reveal the causes of deadlocks. From a quantitative 
point of view, this study revealed that about 30% of exchanged 
messages expressed request of assistance, which represents a 
very significant percentage, which can affect seriously the per-
formance of the learning process. Meanwhile, from the qualita-
tive point of view, the most important assistance requests were 
related with doubts about the use of the computer tool and with 
doubts about links, semantically correct, between underlying 
concepts, which showed the relevance of the development of 
skills to build correct structures.

Derived from these results, we have determined behavior 
patterns associated with the phases (the bootstrap phase, the 
work phase and the goodbye phase) of the learning process 
deefined in this work. Based on these behavior patterns, we can 
count with the following elements to offer the assistance closer 
to the student’s needs: the phase during which the assistance 
request occurs and the kind of help that the student is requested.

As mentioned previously, this work is part of our final objec-
tive towards the development of an automatic computer system 
capable to assist students when they get stuck during the learn-
ing process in on-line mode.

REFERENCES

groups. London: Routledge & Kegan Paul Ltd.


larta: Alta-Omegaf, 383-388.


ligence in Education, 12, 40-62.


Yang, C., & Chang, Y. S. (2010) Assessing the effects of interactive blogging on student attitudes towards peer instruction, learning mo-