9-4-2008

On Incorporating Learning Theories to Simulate a Computer-Supported Collaborative Learning Environment

Nobel Khandaker
University of Nebraska, knobel@cse.unl.edu

Leen-Kiat Soh
University of Nebraska, lsoh2@unl.edu

University of Nebraska–Lincoln, Computer Science and Engineering
Technical Report TR-UNL-CSE-2008-0006
Issued Sept. 4, 2008

On Incorporating Learning Theories to Simulate a Computer-Supported Collaborative Learning Environment

Nobel Khandaker and Leen-Kiat Soh

Computer Science and Engineering
University of Nebraska-Lincoln
256 Avery Hall, Lincoln, NE 66588-0115

(402) 472-6738

{knobel, lksoh}@cse.unl.edu

ABSTRACT

The success of a computer-supported collaborative learning (CSCL) environment in improving student knowledge depends on factors like student attributes (e.g., ability and motivation) and elements of the CSCL environment (e.g., group size and group formation). Better understanding of the individual and combined effects of these factors on the learning outcome of students would help: (1) researchers design better CSCL systems and (2) teachers make better decisions while carrying out CSCL sessions. To aid the researchers and teachers, we have used the observations collected and derived from published theories on individual, peer-based, and collaborative learning to design SimCoL, a multiagent-based tool for simulating the collaborating learning process in a CSCL environment. SimCoL consists of agents that model student collaborative behaviors, teacher instructional decisions, and agent-based learning support in the CSCL environment.

1. INTRODUCTION

Collaborative learning has been used by teachers to facilitate and/or improve student learning through cooperation and interaction in traditional classrooms. Recent advancements in the educational psychology and computer and communication technologies have given rise to computer-supported collaborative learning (CSCL) bringing collaborative learning to students who are not situated in the classroom at the same time. Although CSCL environments are usually intended to increase the overall knowledge of the participating learners, this goal is not easy to achieve as such learning depends significantly on the interactions between the learner and his or her peers in the CSCL environment. Many factors could affect the interactions, including the individual characteristics of the students (e.g., cognitive ability [1], motivation [2], their friendship with others [3]) and the features of the environment (e.g., the group formation method, teacher support for collaboration [6]).

One way to investigate the roles that the various elements of a collaborative learning environment play in collaborative learning is through agent-based simulations. When used to simulate human learning behaviors, these agents can be designed following existing individual and collaborative learning theories with similar uncertainties and limitations. This allows a teacher to simulate "what-if" situations to inform how he or she should carry out a particular collaborative activity given what he or she knows about the students and the learning environment. For example, various student attributes are usually measurable by classroom surveys (e.g., ability and knowledge can be estimated by pre-tests, motivation can be estimated by [7] [8]). So, the teacher can input such collected details of the student attributes into the simulator and estimate the advantages or disadvantages of deploying a CSCL system in his or her classroom. In addition, a simulator for CSCL environment could help researchers in the CSCL community design better CSCL systems such as designing better algorithms for student group formation. A simulation environment for the CSCL system will allow researchers test the effectiveness of their algorithms on large number of students for a long period of time, that they might not be able to do with real classrooms. Finally, many CSCL tools are combined with intelligent agents to improve student learning. These agents are used for forming groups and providing support to the students while they are collaborating [9]. Simulations may allow the researchers to gain a better insight into the usefulness of such agent-based services in CSCL classrooms by studying the usefulness and effectiveness of various agent-based algorithms and services on a larger set of student models before implementing them in the real classroom. So, a simulator for CSCL environment could provide insights into CSCL or collaborative learning process that may take the researchers years and hundreds of students to obtain.

However, existing educational simulation tools, e.g., [10] and [8] and simulation tools designed for group learning scenarios, e.g., [11] and [12] are not fully capable of addressing the complexities of collaborative learning scenarios in a typical CSCL environment. For example, the educational simulation tools [10] and [8] only considered student learning from the teacher and did not accommodate student learning from peers. Furthermore, Spoelstra and Sklar [11] did not consider the cognitive ability of a learner while calculating his or her improvement in knowledge even though, as noted by the researchers working on learning theories, e.g., [1,13] cognitive ability plays a crucial role in determining the individual learning outcome of a student. Therefore, a simulation environment that incorporates the important student-related and classroom-related factors in it could better represent an actual collaborative classroom.

In this paper, we describe SimCoL – a multiagent application for simulating the collaborative learning of a set of students in the CSCL environment. The inspiration source of our paper is CSCL environments that combine research ideas from psychology (especially educational psychology), education, and computer science to create an online collaborative learning environment for students. The primary focus of our research is to build a multiagent simulator in which the agents' behavior, guided by the individual and collaborative learning theories, closely represents the collaborative learning behavior of the students in a CSCL environment. Our primary focus would allow the CSCL researchers and teachers to gain insights into the collaborative learning process and the impact of the various student attributes and teacher-controlled parameters on the learning outcome of the students. The secondary focus of our research is to incorporate an agent architecture in which the agents act as the assistants of the simulated students in the environment. These agents act as assistants to the simulated students and provide services like forming learning groups and supporting their collaboration. This secondary focus would allow CSCL researchers, teachers, and researchers who apply multiagent techniques to CSCL systems to investigate the impact of agent-based services (e.g., agent-based group formation algorithms and agent-based support for students' collaborative learning) on the learning outcome of students. The SimCoL environment consists of an agent that acts as the teacher and agents that act as the students in a CSCL classroom. Furthermore, SimCoL contains an agent framework containing a teacher agent and student agents where the teacher agent helps the teacher in forming groups and the student agents monitor the activities of the students and helps the students form groups. During the simulation of a CSCL classroom session, the teacher: (1) forms student groups using agent-based (VALCAM [14]) and non-agent-based (random and Hete-A [15]) group formation methods, (2) assigns collaborative tasks to the students, and (3) controls environment parameters--such as task difficulty, group size, group formation scheme, and instructor support for collaborative learning. Once the instructor initiates the CSCL session, the students collaborate with each other according to their own models of knowledge, ability, motivation, emotion, and social relationship with others. As a result of this simulated collaboration, the assigned collaborative task is solved by the students and their knowledge on the topic of the task increases. How well the task is solved and how much their knowledge increases are based on: (1) their individual characteristics, (2) the characteristics of their groups, and (3) other teacher-controlled attributes like group formation scheme, and instructor support for collaborative learning. This average improvement of the knowledge of the students due to their collaboration is one way of estimating the success of the CSCL environment design. Thus, the ability of adjusting the various design parameters in response to the improvements in knowledge allows a teacher to: (1) investigate the impact or the appropriateness of a specific CSCL design on different groups of students and (2) identify how to best support collaborative learning given a specific classroom of students.

The rest of the paper is organized as follows. Section 2 presents a set of learning theories and observations based on the relevant research regarding the individual, peer-based, and collaborative learning of students. Section 2 also describes how the teachers support the collaborative learning of students in CSCL classrooms using scaffolding. In Section 3, we use the observations in Section 2 to design the agent that represents the teacher, the agents that represent the students, and the tasks in the SimCoL environment. Section 3 also briefly describes the incorporation of an agent framework (containing a teacher agent and student agents) and a group formation framework (VALCAM [14]) based on that agent framework in SimCoL. Then we describe how the SimCoL environment was realized using Repast—a multiagent simulation tool, in Section 4. Section 5 describes some related work and in Section 7 discuss the conclusions.

2. LEARNING

In this section, we discuss definitions, theories, and empirical observations regarding four different aspects of students' learning processes in a CSCL classroom: (1) individual learning, (2) learning by interacting with a peer, i.e., peer-based learning, (3) collaborative learning in student groups, and (4) teacher support for the collaborative learning process or scaffolding. For individual learning, we discuss the theories that describe how the individual characteristics of a student affect his or her improvement of knowledge due to learning. Second, for peer-based learning, we discuss: (1) the various learning scenarios that are possible when two students are solving a task together and (2) possible situations when peer-based learning may fail. For collaborative learning, we discuss the theories that describe the collaborative learning process for a group of students. We also describe how the collaborative learning of a group of students can be decomposed into series of interactions between two students and how the individual characteristics and social and affective issues impact the interactions between two students. Then for scaffolding, we define the scaffolding process in the CSCL setting and report observations from our review of related research works regarding scaffolding. The reported observations regarding the individual, peer-based, collaborative student learning, and scaffolding are used: (1) in Section 3 as design specifications to build agents that have characteristics similar to students and can simulate the students' interactions with their peers leading to the simulation of collaborative learning process in student groups, (2) in Section 3 to: (a) design the effect of the scaffolding provided by the teacher and (b) the effect of that scaffolding on students.

2.1 Individual Learning

We use the term learning to refer to the improvement in a student's knowledge or expertise on a topic or skill. This learning could be topic-specific, e.g., learning how to solve differential equations; or could be topic-independent, e.g., teamwork or communication skills. According to learning theories [13,16] the four main elements that affect how a person learns are: (1) what the student already knows (knowledge), (2) how able/intelligent the student is (ability), (3) how motivated the student is (motivation), and (4) the emotional state of that student (emotion). The cognitive components that represent these factors are: (1) the crystallized intelligence as accumulated knowledge stored in long-term memory, (2) fluid intelligence as represented by working memory capacity, and (3) motivation as represented by working memory allocation [13], and (4) emotional state [16]. Next, we define these elements in greater detail.

Shell and Brooks [13] use the term knowledge to refer to the accumulated knowledge in a student's long-term memory. The ultimate result of learning would occur as the improvement of the knowledge of the students. Shell and Brooks [13] use ability to represent the cognitive ability or intelligence of a person. They suggest that there are two different parts of ability: fluid intelligence and crystallized intelligence. The fluid intelligence is a fixed entity that deals with general cognitive capacity and crystallized intelligence represents the accumulated knowledge of the student. Furthermore, the fluid intelligence is basically the working memory of a student [13]. However, there is a difference between the absolute working memory capacity a person has and the amount of working memory capacity he or she has available at a particular time for a particular task. The behavior of a person while working on a task and the improvement in his or her knowledge due to learning by working on that task depend upon the amount of working memory that person has available at that time. Further, the amount of working memory available to any person at a time is determined by: (1) his or her existing knowledge for that task, (2) his or her motivation to work on that task, and (3) emotion [13]. Motivation determines why we do what we do [13]. In other words, motivation is the process whereby goal directed behavior is instigated and sustained. Finally, the emotion of a student determines whether the students are feeling happy or sad. Shell and Brooks [13] describes that the prior knowledge and motivation increases the amount of working memory a student has for a task. Furthermore, emotion of a student competes with his or her motivation to occupy the working memory allocation. So, if the student is in a heightened emotional state (too happy or too sad), he or she will not be able to work efficiently to complete the current task since that student's working memory has been occupied by that heightened emotional state. So, based on our discussion of the contribution of individual characteristics of a student on his or her improvement of knowledge, we write our first observations as:

Observation 1: A student's improvement of knowledge of a topic is mainly affected by: (1) his or her existing knowledge, (2) ability, (3) motivation, and (4) emotion.

Observation 2: *The amount of working memory available to a student determines how much he or she can learn.*

Observation 3: *The working memory of a student interacts with his or her prior knowledge and new information (regarding a task) to produce learning and behavior.*

Observation 4: *A student's available working memory for a task can be described as his or her ability for that task.*

Although the aforementioned four components that affect learning are cognitively distinct from one another, they are closely connected with each other and inseparable when they are to be discussed from the point of view of learning of a student [13]. The combined effect of these four components on the learning of a student described by [13] can be summarized as: (1) the prior knowledge stored in the long-term memory interacts with the working memory to produce learning, (2) available amount of working memory limits how much prior knowledge and information can be used/activated at any time, (3) the amount of working memory is determined by motivation, extent of prior knowledge, and emotion, and (4) as knowledge increases, it increases the effective working memory capacity allowing acceleration of future learning processes. Hence, we have the following observation:

Observation 5: *A student's available working memory for a topic is proportional to his or her: (1) knowledge on that topic, and (2) motivation to learn that topic. Furthermore, this available working memory is inversely proportional to the emotional state of that student.*

Observation 6: *As the knowledge of a student on a particular topic increases, his or her learning outcome for that topic would accelerate (if the motivation and emotion stays unchanged) due to increased working memory allocation.*

2.2 Peer-based Learning

When a student is working with his or her peer to solve some assigned task, the student and the peer may learn from each other about that task. The possible learning scenarios between two interacting peers are summarized by [17] as:

Learning by Observation: A student can learn indirectly by observing his or her peer's learning process. This type of learning is more common when the involved students are at the same knowledge level.

Learning by Teaching/Guiding: Learning by teaching occurs when a student learns or refines his or her own knowledge by teaching his or her peer. This type of learning requires that the student teaching the other has prior knowledge about the assigned task.

Learning by Being Taught: This is the simplest type of learning where a student learns when he or she is being taught by his or her peer. Learning by teaching and learning by being taught may work in unison. If a student x is teaching student y, then student x can improve his or her knowledge by teaching, and y can improve his or her knowledge by being taught.

Learning by Reflection/Self-Expression: This type of learning occurs when a student rethinks his or her own solution and analyzes his or her self-thinking process [18], e.g., when a student is explaining his or her solution of a task to his or her peer.

Learning by Apprenticeship: In this type of learning, the expert shows the apprentice how to do a task, watches as the apprentice practices portions of the task, and then turns over more and more responsibility until the apprentice is proficient enough to accomplish the task independently [19]. Note that learning by being taught improves the knowledge or skill of the student who is being taught by someone else. On the contrary, learning by apprenticeship improves the knowledge of the apprentice who is observing and mimicking someone else's behavior.

Learning by Practice: This type of learning occurs when a student applies existing knowledge to solve an assigned problem. This type of learning is very common in situations where two students are jointly solving parts of the assigned task. Notice that in learning by practice, a student improves his knowledge on a topic about which he or she has prior knowledge. However, in learning by apprenticeship or learning by being taught, a student learns something about which he or she does not have any prior knowledge.

Learning by Discussion: This type of learning occurs when two students discuss the solution of a task with each other. Notice that this type of learning is basically a sequence of Learning by Observation, Learning by Teaching, Learning by being Taught, Learning by Reflection/Self-Expression, and Learning by Practice except that the roles of the students are dynamic in Learning by Discussion.

From our summarization of the peer-based learning, we observe that the prior knowledge of the participating students plays an important role in deciding what type of learning scenarios may occur. For example, learning by teaching (and learning by being taught) is more common among two students where one student with prior knowledge teaches his or her peer who has less prior knowledge. We can summarize the possible peer-based learning scenarios according to the prior knowledge of the students as observations in Table 1.

Table 1. Possible Learning Scenarios among Peers

Observation	Student's Knowledge	Peer's Knowledge	Learning by

7	High	High	Observation, Reflection, Practice and Discussion
8	High	Low	Observation, Reflection, Practice and Discussion, Teaching, and Being Taught
9	Low	High	Observation, Reflection, Practice and Discussion, Teaching, and Being Taught
10	Low	Low	Observation

Furthermore, the difference between two interacting students' prior knowledge about how to solve a certain task can hinder their learning. This effect is described in Vygotsky's zone of proximal development (ZPD) theory [20]. Vygotsky describes that learning is most likely to occur, when the teacher and the student are able to discuss the subject matter in a manner that is understandable to the student. Therefore, it may be frustrating and difficult for two students to learn from each other if the amount of prior knowledge they have on a topic is very different from each other [11]. For example, if a student is trying to prove that a certain problem is NP-complete, it is unlikely that he or she would benefit by discussing the solution approach with someone who does not know anything about algorithms. This gives us our next observation:

Observation 11: Two students may learn about a topic from their interactions (Table 1) when the amounts of prior knowledge they have are not too different from one another.

2.3 Collaborative Learning

The term ”collaborative learning” is an instruction method in which students at various performance levels work together in small groups toward a common goal [5,21]. In this sub- section, we discuss how a group of students can interact and improve their knowledge. First, we describe the collaborative knowledge building process of a group of students. Then we report empirical observations of CSCL researchers to define the collaborative learning process of a group of students as a series of interaction or discourse between group members. Finally we discuss the observations of CSCL researchers that describe the effect of social and affective issues on the learning outcome of a group of students.

Collaborative Learning as a Cyclic Process: The goal of collaborative learning is to allow the group members to increase their knowledge on a specific topic. Stahl [22] describes the collaborative knowledge building process for a group of students through the interactions of those students. The collaborative knowledge building process as described by Stahl [22] can be summarized using the following observations:

Observation 12: *The collaborative knowledge building is a cyclic process that feeds on itself and converges exponentially faster.*

Observation 13: *This collaborative knowledge building cycle is a hermeneutic cycle, meaning, ”one can only interpret what one already has an interpretation of”.*

Observation 14: *Individual knowledge of a student is gained from collaborative knowledge of his or her group members through interaction. That collaborative knowledge is in turn produced by individual knowledge of the interacting group members.*

Collaborative Learning through Interaction/Discourse: We have already discussed that a student's knowledge can improve due to that student's collaboration with his or her peer. Now we describe how that collaboration can occur between two collaborating students. Kreijns [23] describe the interaction between students as the key to collaboration among group members. Furthermore, empirical evidence suggests that collaborative learning in a CSCL occurs from the exchange of dialogues among the students [24,23]. So, from these observations, we can describe our next observation as:

Observation 15: *The collaboration among the members of a group of students occurs due to their interaction/discourse with each other.*

The discourse/interaction among the students may be of different shapes and forms (e.g., task-dependent, or task-independent). Zumbach [25] describes a collection of dyadic (between two students) interactions for a group of students which were reported by researchers in the CSCL community. An example of interactions mentioned in [25] is: (a) student a proposes a solution for the assigned task, (b) student b accepts or proposes another solution to the task. We write our observation regarding the interaction between two students as:

Observation 16: *The compilation of discourse/interaction patterns presented by Zumbach et al.* [25] *describes a typical dyadic (between two students) learning scenario in terms of a chain of action-reaction patterns. In* [25]*, the authors describe nine different possible utterances (e.g., proposal of a solution) to start collaboration. Each of these possible starting utterances has one or more possible replies (e.g., query or challenge the proposal). Finally, each of those replies has one or more possible reactions (e.g., modify the solution proposal) from the initiator.*

Affective and Social Issues: Zumbach et al. in [25] describe typical interaction scenarios in a CSCL environment. However, how many and of what quality these discussions/interactions may occur, depends on the affective state of a student [2] and his or her social relationship with other students in the group. Furthermore, the individual affective state of a student and his or her social relationship with others depends on a variety of aspects of a typical CSCL environment. These factors include: (1) teacher's feedback, (2) the nature of the task, and (3) length of the collaborative session, etc. Next, we describe those issues in greater detail as described in [2] as cited in [26].

- **Social relationship between partners**: The social relationships among a group of students denote their mutual respect for each other and their willingness to work together. Issroff and Del Soldato [2] argue that social affinity has a significant effect on the nature and effectiveness of collaborative interaction since the people who are used to working together have established ways of negotiating their individual and common goals. Group members who are new to each other, on the other hand, have to negotiate the rules of interaction which could be demotivating for some students. In addition, Jones and Issroff [26] and Vass [27] report that, students who are friends have established ways of working which are implicitly understood rather than explicitly discussed. Furthermore, friends typically have a better grasp of each other's ideas and state of knowledge which is crucial for successful collaboration [26]. The effect of social relationship on the collaborative learning performance of a group has also been discussed by [28,23] where it is mentioned that social relationships contribute to common understanding, an orientation towards cooperation, and the desire to remain as a group. The social relationships among students also change as they collaborate. As reported in [9], the students form their view of other students due to the type and extent of collaboration they receive from their peers.

- **Time**: The quantity (how many) and quality (how much it improves student knowledge) of interaction and/or collaboration varies over the length of the collaborative sessions due to two factors. First, students may develop friendship with their group members due to their collaboration which would motivate them to have more interaction with them and vice versa. Second, the students' motivation and emotion may change (due to interaction with other students or due to factors external to the classroom e.g., illness) over time. Those changes in a student's characteristics would then improve or lessen his or her interactions with the group members.
- **Group Member's Motivation**: Clear and Kassabova [29] report that in CSCL classrooms it is common to have students whose motivation is affected by the motivation of other group members. When the other group members are motivated to learn and to collaborate, it increases the motivation of a student who had low motivation when he or she joined the group. On the other hand, when a student joins a group with high motivation and finds the peers to be not too motivated, it may reduce his or her motivation.

We derive from the above the following observations:

Observation 17: *Good social relationship and/or friendship improve the quantity and quality of interaction among a group of students.*

Observation 18: *The quantity and quality (i.e., learning outcome) of interactions among a group of students varies over time due to factors internal and external to the classroom environment. Improvement in social relationship among the members of a group improves the quality of collaborations among them. On the other hand, when a student group member experiences distracting factors, that experience reduces the quality of his or her collaboration with the other group members.*

Observation 19: *Motivation of the group members' may have positive or negative effect on the motivation of a student. If the group members are motivated, it may increase that student's motivation and if the group members are not motivated, it reduces that student's motivation.*

Observation 20: *Social relationship between a student and his or her peer (as perceived by the student) may change according to the frequency and extent of collaboration (e.g., how many times did my peer helped me). If the peer helps the student complete the assigned task by collaborating with him or her, the social relationship improves, otherwise, the social relationship deteriorates.*

2.4 Scaffolding

Bruner [30] and Cazden [31] define scaffolding as the act of providing assistance to a child so that he or she is able to carry out a task (e.g., solve a problem) that he or she cannot do by herself. As cited in [6], Greenfield [32] (p. 118) describes the scaffolding process with five characteristics: (1) it provides support, (2) it functions as a tool, (3) it extends the range of the worker, (4) it allows the worker to accomplish a task not otherwise possible, and (5) it is used to selectively aid the worker where needed. The original notion of scaffolding was developed to address situations where an expert (e.g., a teacher or an adult) would help an inexperienced learner by providing him or her exactly the type of help needed to complete a task. However, over time, the concept of scaffolding has been introduced into traditional classrooms to aid learners to achieve difficult learning objectives and complete difficult tasks. Puntambekar and Hubscher [6] describe that the notion of scaffolding has been increasingly incorporated in the classroom teaching where the scaffolding is provided by various types of software tools. According to Puntambekar and Hubscher [6], the tools and software that are used to provide scaffolding usually are used to: (1) offer structure and support for completing a task and (2) promote peer interactions to enable peers to support each other's learning. In the first type of scaffolding, the students are provided information about how to better approach to solve the task that they are having difficulty with. In the second type of scaffolding, the peer support of a student is enhanced in the hope that those peers would provide guidance and information for that student to help him or her solve that task. Like traditional classrooms, the use of software to provide scaffolding has been embraced by the CSCL community too. Researchers in the CSCL community are now utilizing scaffolding in the form of incorporating structure of learning activities (e.g., [33]) and improving peer support (e.g. [34]). The design and implementation of the scaffolding process in the CSCL environments usually require additional cost and effort [33]. However, the scaffolding process in a CSCL environment can be used to improve the learning of a large number of students which is difficult to do by the instructor alone in a classroom. On the other hand, it is easier for the instructor to determine the need and level of understanding of a learner and provide learner-specific scaffolding than a software that is designed to provide scaffolding. As CSCL researchers (e.g., [6]) note that due to being in different zones of proximal development, the learners benefit most when the scaffolding is targeted toward their zone of development. So, one of the recommendations provided to the CSCL practitioners is to customize the scaffolding to specific learners' needs. So, from our discussions of the scaffolding process, we write the following observations:

Observation 21: *Scaffolding in the CSCL environment can be provided by: (1) providing structure and support for completing tasks and (2) improving of peer support.*

Observation 22: *Scaffolding in the CSCL environment may be used to improve the knowledge of the learners regarding the assigned task.*

Observation 23: *Learners in a CSCL environment benefit more when the provided scaffolding is targeted to their zone of proximal development.*

3. SIMCOL ENVIRONMENT

The SimCoL environment E represents a CSCL environment where the teacher forms student groups and assigns a set of tasks and the students solve those tasks collaboratively to improve their knowledge about some topic. The SimCoL environment is defined as a 5-tuple: $E = \langle T, I, H, SA, ta \rangle$. Where T is a set of tasks, I is an agent who acts as the teacher, and $H = \{h_1, \dots, h_{ns}\}$ is a set of agents who represent the students in a collaborative classroom environment. Furthermore, $SA = \{sa_1, \dots, sa_n\}$ is a set of student agents and ta is the teacher agent. Each student agent in SimCoL is assigned to a student and the teacher agent is assigned to the instructor. In this section, we first define the tasks T. Then, based on the observations presented in Section 2, we describe the attributes and the behavior of agents H who represent the students in SimCoL. Furthermore, we describe how the teacher I forms student groups and carries out CSCL classroom sessions in the SimCoL environment using a set of simulation steps. Then we describe the roles of the student agents and the teacher agent and briefly discuss how the student and teacher agents work together to form student groups using the VALCAM [14] algorithm. Finally, we describe the collaboration process of the students H in a group in SimCoL using a set of simulation steps and discuss how the students' attributes change due to collaboration.

3.1 Task

The tasks in SimCoL represent the problems and exercises that are solved by the students in a CSCL environment. The set of tasks is denoted by, $T = \{T_1, \dots, T_{nt}\}$ where,

$$T_j = \langle ct_j, di_j, ts_j, tl_j, \overrightarrow{sq_j} \rangle \quad (1)$$

In (1):

ct_j denotes the concept of the task. This concept represents the subjective knowledge required to solve the task.

$di_j \in [ll_{di}, ul_{di}]$ (with $ll_{di}, ul_{di} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) is the difficulty of the task as determined by the teacher.

$ts_j = \{ts_{j1}, \dots, ts_{jn}\}$ is the set of sub-tasks in T_j.

$tl_j \in [ll_{tl}, ul_{tl}]$ (with $ll_{tl}, ul_{tl} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) is the time limit within which the task is to be completed.

$\overrightarrow{sq_j} = \{sq_{j,t} | for\ time\ t\}$, where $sq_{j,t} \in [ll_{sq}, ul_{sq}]$ (with $ll_{sq}, ul_{sq} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) is a vector representing the students' (i.e., students working on the task) view of the solution quality of the assigned task T_j at time t.

To illustrate how the above tuple factors into a task, take T_j as an instance: "Write an essay describing the pros and cons of Vickrey [36] auction protocol". The concept ct_j can then be described as: "Vickrey auction protocol." The difficulty di_j would be determined by the teacher based on the students' knowledge and experience on ct_j. Furthermore, the subtasks ts_j could be: 1) write introduction, 2) write the pros of Vickrey Auction, 3) write the cons of Vickrey Auction, 4) write the conclusion, 5) proof-read, and (6) check the logical flow of the essay. The time limit tl_j could be set by the teacher (e.g., 7 days). The solution quality $\overrightarrow{sq_j}$ would be a vector of real values that represent the quality of the solution from the perspective of the students working on task T_j at time t and change over time.

3.2 Student

We represent the model $hm_{i,t}$ of each student $h_i \in H$ in SimCoL by a 6-tuple:

$$hm_{i,t} = \langle K_{i,t}, AB_{i,t}, MO_{i,t}, EM_{i,t}, SR_{i,t}, TSQ_{i,t} \rangle \quad (2)$$

In (2):

$K_{i,t} = \{(ct_j, ex_{i,j,t}) | \forall T_j \in T\}$ is the knowledge of student h_i at time t with ct_j representing the concept of T_j and $ex_{i,j,t} \in [ll_{ex}, ul_{ex}]$ (with $ll_{ex}, ul_{ex} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) representing the expertise, i.e., the amount of knowledge the student has about the concept. The goal of student collaboration is to increase the value of this expertise.

$AB_{i,t} = \{ab_{i,j,t} | \forall T_j \in T\}$ where $ab_{i,j,t} \in [ll_{ab}, ul_{ab}]$ (with $ll_{ab}, ul_{ab} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) is the ability of h_i at time t for task T_j.

$MO_{i,t} \in [ll_{mo}, ul_{mo}]$ (with $ll_{mo}, ul_{mo} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) is the motivation of h_i at time t.

$EM_{i,t} \in [ll_{em}, ul_{em}]$ (with $ll_{em}, ul_{em} \in \mathbb{R}$, denoting the lower- and upper-bounds, respectively) is the emotional state of student h_i at time t.

$SR_{i,t} = \{sr_{i,k,t} | h_k \in H - h_i\}$ where $sr_{i,k,t} \in [ll_{sr}, ul_{sr}]$ (with $ll_{sr}, ul_{sr} \in \mathbb{R}$, denoting the lower- and upper-bounds respectively) is the social relationship between h_i and h_k at time t as perceived by h_i.

$TSQ_{i,j} \in [ll_{tsq}, ul_{tsq}]$ (with $ll_{tsq}, ul_{tsq} \in \mathbb{R}$, denoting the lower- and upper-bounds respectively) denotes the target solution quality of the task T_j of h_i at time t.

Notice that, we have included $K_{i,t}$, $AB_{i,t}$, $MO_{i,t}$, and $EM_{i,t}$ in the student model according to Observation 1 and included $SR_{i,t}$ according to Observations 17 and 18. In addition, combining Observations 4 and 5, we assume that the ability of a student is related to his or her knowledge, motivation, and emotion in the following way:

$$\begin{aligned} AB_{i,j,t} = AB(hm_{i,t}) \propto\ & w_{abx} \cdot ex_{i,j,t} + \\ & + w_{abm} \cdot MO_{i,t} - w_{abe} \cdot |EM_{i,t}| \end{aligned} \quad (3)$$

where w_{abx}, w_{abm}, and w_{abe} are weights. So, according to (3), the ability of a student for a particular task at any time is proportional to the sum of his or her expertise on the concept of that task and motivation minus the absolute value of his or her emotional state. So, if a student has high expertise and motivation and has stable (or non-elevated, i.e., not too happy or not too sad) emotional state, he or she will have a higher ability and vice versa. We also define the target solution quality of a student as,

$$TSQ_{i,j,t} \propto AB_{i,t} \quad (4)$$

So, a student's target of the quality of the solution of the assigned task is proportional to his or her ability for that task. According to Observations 3 and 4, the ability of a student determines how much of his or her existing knowledge can be activated to produce behavior (i.e., effort to solve the task) and learning. Therefore, given the same time limit tl_j for a task T_j, a student with higher ability would be able to solve the assigned task better than a student with lower ability. So, we assume that the students have targets of the final solution quality according to their own abilities. A student with high ability (i.e., high expertise, motivation and stable emotional state) for an assigned task would aim to complete the assigned task very well, i.e., high target task solution quality, and vice versa.

3.3 Teacher

The teacher I in SimCoL acts as the coordinator of the CSCL sessions. The teacher delivers instructions, forms groups, and assigns collaborative tasks. In SimCoL, we have implemented three different group formation methods: random, Hete-A, and VALCAM group formation method.

In random group formation, the teacher forms groups by randomly selecting students and assigning them to different groups. Random group formation has been used by the researchers in the CSCL community: (1) to evaluate the performance of group formation methods by comparing the learning outcome of the students collaborating in the randomly formed groups and groups formed by other methods (e.g., [9]) and (2) as a low complexity algorithm for forming heterogeneous groups [37].

In SimCoL, the teacher uses the Hete-A algorithm [15] to form heterogeneous groups. The Hete-A algorithm forms student groups using the Matrix-Hete which is a two-dimensional matrix of student characteristics. The row and column of the matrix represent the different values of the two characteristics and the value in each cell is the number of students whose characteristic values are equal to the values in the row and column. The Hete-A algorithm works in the following way. First, the cell with the highest value is chosen. One student from this cell is randomly chosen and put into a group and the value of the cell is decreased by one. Then the row and column to which the selected student belonged to is excluded and the next student member is selected. This procedure is repeated until the first student group is formed. This group selection process is then repeated until all student groups are formed. If there are multiple cells that have the same highest value then the cell with the lowest parameter ideal distance (Euclidean distance between the cell having the highest value in the previous step and the cells having the same highest value specified in the current step) is chosen. If at some point, the current group cannot be completed although there are still remaining students because all rows and columns have been excluded, all rows and columns are recovered with their updated values and the process continues.

Next, in Table 2, we discuss how the teacher carries out the CSCL session through a set of simulation steps. First, the teacher initializes the tasks, chooses the group formation scheme (Step 1), chooses how often scaffolding should be provided, and how many groups would receive scaffolding. Then, for each initialized task, the teacher conducts a collaborative session. During initialization (Step 2a), the teacher initializes the time (Step 2a(i)) and the student groups (Step 2a(ii)), announces the task to all students (Step 2a(iii)), and chooses the minimum group size (Step 2a(iv)). Then, if the group formation is random, the teacher forms student groups by either randomly assigning a student to a group (Step 2b(i)–(ii)) or using the Hete-A algorithm [15] (Step 2c) to form student groups. Once the groups are formed, the teacher announces the start of the collaborative session to all students (Step 2e). Then after every scaffolding period until the collaborative session is over, the teacher sorts the groups according to their current solution quality of the task (Step 2f(i)-a(1)) and then selects the groups who have the lowest solution quality. Those selected groups are then provided scaffolding (Step 2f(i)-a(2)). Finally, the teacher announces the end of the collaborative session when the time limit for the current task is over (Step 2f(ii)).

Table 2. Simulation Steps of Teacher

Simulation Steps of Teacher I
1. Initialization: $T \leftarrow \{T_1, \ldots, T_{nt}\}$, $G_{fs} \leftarrow$group formation scheme, $t_{sc} \leftarrow$scaffolding period, $n_{sc} \leftarrow \lvert G_{sc} \rvert \cdot r_{sc}$
2. For all tasks $T_j \in T$, do,
a. Initialize collaborative Session s_j:
i. $t \leftarrow 0$, $G \leftarrow G_1, \ldots, G_m$,
ii. Announce task T_j to students H,
iii. $n_g \leftarrow \lfloor H_m / m \rfloor$
b. If $G_{fs} = R_n$ Then
i. $H_t \leftarrow H$, $k \leftarrow 0$
ii. While $\lvert H_t \rvert > 0$, do,
1. Choose kth group G_k from G,
2. Randomly choose $h_k \in H_t$,
3. $H_t \leftarrow H_t - h_k$,
4. $G_k \leftarrow G_k \cup h_k$,
5. $k \leftarrow (k+1) \mathrm{mod} \lvert G \rvert$
c. Else If $G_{fs} = H_a$ Then
Use Hete-A algorithm [15] to form groups
d. Else If $G_{fs} = H_a$ Then
Use VALCAM algorithm [14] (Section 3.4) to form groups
e. Announce start of collaborative Session s_j to H
f. While (true)
i. If $t < tl_j$
a. If $\mathrm{mod}(t, t_{sc}) = 0$
1. Sort (ASC) G according to $tsq_{j,t}$
2. For $i \leftarrow 0$ to n_{sc}
Provide scaffolding to G_i
b. $t \leftarrow t + 1$
ii. Else Announce end of collaborative Session s_j to H

3.4 Agents

The student and teacher agents have been incorporated in SimCoL to implement various agent-based coalition formation algorithms. Each student agent in SimCoL is assigned to a student and it monitors the change in that assigned student's: (1) expertise gain and (2) social relationship with other students. The teacher agent is assigned to the instructor to: (1) assign and monitor student collaborative performances and assign them virtual currency according to that performance and (2) communicate with the student agents to form groups using VALCAM [14] – an auction-based group formation algorithm.

In VALCAM, the teacher agent hosts iterative auctions and the student agents bid in those auctions to buy membership to the student groups. The details of VALCAM can be found in [14]. However, a brief description is as follows: in the SimCoL's adaptation of VALCAM, SA is the set of student agents and n_g is the number of student groups T_j is the current task assigned, p is the selected auction protocol e.g., Vickrey[14].

VALCAM-S (for teacher agent ta)

1. Initialize (create a set of n_g groups G)
2. Choose first members for each group $g \in G$ (select better-performing users as first members)
3. Start the auction according to p for student agents in SA. For each group g in G, do,
   a. Accept bids from the unassigned users
   b. Assign the highest bidder to g. Collect second highest bid amount from highest bidder
4. After completing j, assign individual payoff (proportional to number of successful individual collaborations) and group payoffs (proportional to the task solution quality achieved by the group) to SA

VALCAM-U (for student agent s_i)

1. Initialize (estimate and announce the student's competence for the upcoming task)
2. For each round of bidding for group g, bid with an amount proportional to the average of average compatibility between s_i and the members of g and the average competence of the students in g.

In VALCAM algorithm, the teacher agent initializes each student group with a seed student chosen by sorting all students according to their average expertise on the upcoming tasks (T_j) (Step 1, 2 VALCAM-S). Then the teacher agent invites all student agents to bid for each of the groups until all students are assigned to some groups (Step 3(a) VALCAM-S). For each round of bidding, the student agents calculate the bid for the auctioned group by averaging the competence and compatibility of the students in that group and its assigned user. The teacher agent then collects the second highest bid amount from the highest bidder and assigns the highest bidder to the auctioned group (Step 3b VALCAM). Once the groups have completed the assigned task, the teacher agent assigns virtual currency to the students in two-parts: the individual payoff and group payoff. The individual payoff is proportional to the number of successful collaborations completed by the student (as tracked by the student agent) and the group payoff is proportional to the final solution quality (as monitored by the teacher agent) of the solution prepared by a student's group.

3.5 Collaboration and Scaffolding

As described in Observations 15 and 16, the collaborative behavior of a group of students can be broken down to a series of dyadic (i.e., between two students) interactions. So, in SimCoL, we simulate the collaborative behavior (i.e., collaboration to solve the assigned task and to improve expertise) of a group of students using a series of dyadic interactions among the group members. Here, we describe how the interactions between two students are simulated in SimCoL. First, we define the following:

- $MSU(MO_{i,t}, G_m)$ updates the motivation of student $h_i \in G_m$ according to his or her group members' motivations. We define the following:

$$MSU(MO_{i,t}, G_m) = [w_{omo}\, MO_{i,t} + w_{gmo} \sum_{k \in G_m - h_i} MO_{k,t} / (|G_m| - 1)] \quad (5)$$

  where w_{omo} and w_{gmo} are weights. Based on Observation 19, in (5), a student's motivation level is adjusted by calculating a weighted sum of its motivation and the average motivation of other group members. If the average motivation of other group members is higher than the student's motivation, the student's motivation level is increased, and vice versa.

- $CP(h_i, h_k, T_j, t)$ calculates the probability of h_i accepting a collaboration request regarding task T_j from h_k at time t. We define,

$$CP(hm_{i,t}, hm_{k,t}, T_{j,t}) = w_{csr} \cdot sr_{i,k,t} + w_{csq} \cdot (sq_{m,t} - TSQ_{i,j,t}) \quad (6)$$

  where $\boldsymbol{w_{csr}}$ and $\boldsymbol{w_{csq}}$ are weights. So, according to (6), the probability of a student accepting the collaboration request from his or her group member at time **t** is equal to the weighted sum of: (1) the social relationship between that student and the group member and (2) the difference between the quality of the solution and that student's target quality of the solution, all at time $\boldsymbol{t}$. So, a student is more likely to collaborate when the social relationship between the student and the peer is good (Observation 17) and the quality of the solution is below the student's target (i.e., the student thinks the task needs more work).

- $cc_{i,k,t,j} = \{act_{i,k,t,j}, rct_{i,k,t,j}, lct_{i,k,t,j}\}$ denotes a collaboration cycle (similar to the action-reaction patterns mentioned in Observation 16) completed by student h_i with student h_k at time t for task T_j. Here, $act_{i,k,t,j}$ denotes an utterance of action, $rct_{i,k,t,j}$ denotes an utterance of reaction in reply to the action $act_{i,k,t,j}$, and $lct_{i,k,t,j}$ denotes the reaction in reply to the reaction $rct_{i,k,t,j}$.

- $ci_{i,k,t,j}$ denotes a collaboration cycle initiated by student h_i but declined by student h_k at time t for task T_j.

- $CC_{i,k,t,j} = \{cc_{i,k,t,j}\}$ denotes the set of collaboration cycles completed by student h_i with student h_k at time t for task T_j.

- $CI_{i,k,t,j} = \{ci_{i,k,t,j}\}$ denotes the set of collaboration cycles initiated by student h_i with student h_k at time t but was declined by h_k for task T_j.

- $CY_{i,k,t,j} = \{CC_{i,k,t,j}, \dots, CC_{k,i,t,j}, \dots, CI_{k,i,t,j}\}$ denotes the set of all collaboration cycles between students h_i and h_k regarding a task T_j.

- $SQU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j})$ denotes the function that calculates the improvement of the solution quality of a task when two students have collaborated in a cycle $cc_{i,k,t,j}$ to solve T_j. This function is defined as,

$$SQU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j}) = 0 \quad \mathit{If} \quad p_{sq} \geq \kappa_{sq} \qquad \propto (ab_{i,j,t} + ab_{k,j,t})/di_j \ \mathit{Otherwise} \quad (7)$$

  where $\kappa_{sq} \in [ll_{sq}, ul_{sq}]$ (with $ll_{sq},\ ul_{sq} \in \mathbb{R}$) denotes the solution quality update probability threshold and p_{sq} is a random number that is drawn from a uniform random distribution and p_{sq} is within the range $[ll_{sq}, ul_{sq}]$. These two values together determine, regardless of the attributes of the students h_i and h_k, whether the solution quality of the task improves due to the collaboration between students h_i and h_k. For instance, if the value of the threshold κ_{sq} is high, then value of p_{sq} would be smaller than κ_{sq} most of the time it is drawn (due to the uniform nature of the distribution). As a result, most of the time, the collaborations among the students would fail to improve the quality of the solution of the assigned task. On the other hand, if the value of the threshold κ_{sq} is low, then p_{sq} would be smaller than κ_{sq} most of the time it is drawn (due to the uniform nature of the distribution). As a result, the most of the time, the collaborations among the students would succeed to improve the quality of the solution of the task. The use of the update probability threshold and the random number accommodates the fact that even when two students with high abilities are collaborating to solve the assigned task, their collaborative effort may not improve the quality of the solution of the task due to some unforeseen reason (e.g., the students did not understand the requirement of the task). If the quality of solution increases due to collaboration, that increase is proportional to the sum of the abilities of the two students divided by the difficulty of the task. According to Observations 3 and 4, the ability of a student determines how much of his or her existing knowledge can be activated to produce behavior (i.e., effort to solve the task) and learning. Therefore, the higher the ability (i.e., higher expertise and motivation and stable emotional state) of the two students are, the more they will be able to improve the quality of the solution of the task during a collaboration cycle. Furthermore, the higher the difficulty (as assessed by the teacher) of the task is, the less the improvement of the solution will be due to the completion of a collaboration cycle by two students.

- $HEU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j})$ is a function that calculates the improvement in the expertise of h_i for concept ct_j of T_j due to the collaboration cycle $cc_{i,k,t,j}$. We define,

$$HEU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j}) = 0 \; If \; DE(hm_i, hm_k, T_j) > \kappa_{zone} \propto [w_{hab}\, ab_{i,j,t} + w_{hde}\, DE(hm_{i,t}, hm_{k,t}, T_j)] \quad \text{Otherwise} \tag{8}$$

Here,

$$DE(hm_{i,t}, hm_{k,t}, T_j) = |ex_{i,j,t} - ex_{k,j,t}| \tag{9}$$

Here, κ_{zone} is the zone of proximal development constant and w_{hab} and w_{hde} are weights. Furthermore, Eq. 9 calculates the absolute difference between the expertise of two students. So, the improvement in the expertise of a student for a particular task is calculated by Eq. 8, where:

- According to Observation 14, the expertise of a student can be improved as a result of interaction or collaboration with a peer.
- However, if the difference between the expertise of the collaborators is too large (i.e., larger than the zone of proximal development constant κ_{zone}), then the collaboration does not increase the knowledge of either student. This corresponds to Observation 11.
- The increase in expertise (due to collaboration with a student h_k) of a student h_i is proportional to the weighted sum of h_i's ability and the difference between the expertise of h_i and h_k. The higher the ability of h_i is and the higher the difference between the expertise of h_i and h_k is (as long as the difference is within the proximal development zone), the higher the increase in expertise is. The use of the ability of the student is motivated by Observations 3 and 4. Furthermore, the use of the difference of expertise function DE in Eq. 8 addresses Observations 7—10 that say: (1) the possible learning scenarios between two peers largely depend on their prior knowledge (i.e., expertise) and (2) most of the learning scenarios occur when the expertise values of the students are not the same, i.e., one high and one low.
- $SRU(CY_{i,k,t,j})$ is the social relationship update function defined as:

$$SRU(CY_{i,k,t,j}) \propto [|CC_{i,k,t}| - |CI_{i,k,t,j}|]/[|CC_{i,k,t,j}| + |CI_{i,k,t,j}|] \tag{10}$$

- Notice that according to Observation 20, the social relationship between two students, which is simply a normalized ratio, is updated according to the failure and success of their collaborations. The more successful collaborations they have during the session, the better their social relationship becomes.
- $SEU(hm_{i,t}, ct_j, SO_j)$ is a function that calculates the improvement in the expertise of h_i for concept ct_j of T_j due to the scaffolding object SO_j. We define,

$$SEU(hm_{i,t}, ct_j, SO_{k,j}) \propto 1/(1 + |ex_{i,j,t} - sol_j|) \tag{11}$$

$If \; p_{sca} \geq \kappa_{sc}$ and $= 0$ Otherwise

Notice that according to Observation 21 and 22 as formulated in (11), the SEU function denotes the improvement in expertise of the students due to scaffolding object $SO_j = \langle ct_j, sol_j, csc_j \rangle$ that provides structure and support for the students. Here, $ct_j \in T_j$ denotes the task concept the scaffolding is targeted for, sol_j denotes the level of expertise for the student the scaffolding is designed for, and csc_j denotes the cost of the scaffolding. Here, the cost of the scaffolding denotes the time and effort required to prepare this scaffolding object. According to Observation 23, in Eq. 11, the improvement in the expertise of the students is inversely proportional to the difference of the expertise of the student and the zone of the target students. Furthermore, the occurrence of this improvement depends on the scaffolding acceptance probability value p_{sca}. This probability value is drawn from a uniform distribution and denotes the probability that a student will accept the scaffolding provided to him or her. The improvement in expertise of the student occurs when this probability value is greater than the threshold κ_{sc}. Finally, according to Observation 23, the improvement in expertise of the student is high when the expertise of the student receiving the scaffolding is same as the expertise of the students for who the scaffolding was designed. Otherwise, the effect of the scaffolding is low.

Table 3 shows the simulation steps of a student in SimCoL. During the initialization, the student receives its group assignment and the task (Step 1) from the teacher (Step 2a(iii) in Table 2). After receiving the group assignment, the student updates its own motivation according to other group member's motivations using Eq. 5 (Step 2) and updates its ability (Step 3) accordingly. Then, once the student receives the announcement of the start of the collaborative session from the teacher (Step 2d in Table 2), the student starts the collaborative session (Step 4). During the session, the student keeps track of all the group members it interacts with using a list of collaborators (Step 5a) and checks whether the solution quality of the task is greater than or equal to its own target solution quality (Step 5b). If the current solution quality is less than its own target solution quality, the student sends a collaboration request to one of the group members (Step 5b(i)). If the group member agrees (Step 5b(ii)), then the student completes and stores the collaboration cycle (Step 5b(ii)(1)-(2)), updates the solution quality of its group's assigned task using Eq. 7 (Step 5b(ii)(3)), and updates its own expertise using Eq. 8 (Step 5b(ii)(4)-(5)). If the group member declines the request (Step 5b(iii)), then the student stores the declined collaboration request (Step 5b(iii)(1)). Meanwhile, if the student receives a request for collaboration from any of its group members (Step 5c) and if the probability of collaboration with that student (Eq. 6) is higher than the collaboration threshold, the student completes the collaboration cycle (Step 5c(i)(1)), stores the completed collaboration cycle (Step 5c(i)(2)), and updates its own expertise (Step 5c(i)(3)-(4)). If the probability of collaboration is smaller than the collaboration threshold, then the student declines the collaboration request (Step 5c(ii)(1)) and stores the failed collaboration cycle (Step 5c(ii)(2)). Correspondingly, the student stores the group member who requested the collaboration in its list of collaborators (Step 5c(iii)). During the collaborative session, if the student receives scaffolding from the teacher (Step 5(d)) in the form of a scaffolding object, it updates its expertise using Eq. 11. Finally, when the collaborative session ends, (i.e., announced by the teacher in Step 2e(ii) in Table 2) the student updates its own view of its social relationship with all its collaborators (Step 6(i)-(iii)) using Eq. 10.

Table 3. Simulation Steps of Student

Simulation Steps of Student h_i

1. G_m ←assigned group by the teacher, T_j ←assigned task
2. $MO_{i,t} \leftarrow MSU(MO_{i,t}, G_m)$
3. $AB_{i,t} \leftarrow AB(hm_{i,t})$
4. Wait for the start of collaborative session s_j
5. Until collaborative Session s_j is over, do,
   a. $H_c \leftarrow \phi$
   b. If $sq_{j,t} < esq_{i,j,t}$ Then
      i. Propose collaboration to randomly chosen student $h_k \in G_m - h_i$
      ii. If h_k agrees Then
         1. Complete collaboration cycle $cc_{i,k,t,j}$
         2. $CC_{i,k,t,j} \leftarrow CC_{i,k,t,j} \cup cc_{i,k,t,j}$
         3. $sq_{j,t} \leftarrow SQU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j})$
         4. $\Delta ex_{i,j,t} \leftarrow HEU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j})$
         5. $ex_{i,j,t} \leftarrow min(ex_{i,j,t} + \Delta ex_{i,j,t}, ul_{ex})$
      iii. Else
         $CI_{i,k,t,j} \leftarrow CI_{i,k,t,j} \cup ci_{i,k,t,j}$
      iv. $H_c \leftarrow H_c \cup h_k$
   c. If received collaboration request from h_k Then
      i. If $CP(hm_{i,t}, hm_{k,t}, T_{j,t}) > \kappa_{ch}$ Then
         1. Complete collaboration cycle $cc_{i,k,t,j}$
         2. $CC_{k,i,t,j} \leftarrow CC_{k,i,t,j} \cup cc_{k,i,t,j}$
         3. $\Delta ex_{i,j,t} \leftarrow HEU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j})$
         4. $ex_{i,j,t} \leftarrow min(ex_{i,j,t} + \Delta ex_{i,j,t}, ul_{ex})$
      ii. Else
         1. Decline collaboration request from h_k
         2. $CI_{k,i,t,j} \leftarrow CI_{k,i,t,j} \cup ci_{k,i,t,j}$
      iii. $H_c \leftarrow H_c \cup h_k$
   d. If received scaffolding SO_j Then
      1. $\Delta ex_{i,j,t} \leftarrow SEU(hm_{i,t}, ct_j, SO_j)$
      2. $ex_{i,j,t} \leftarrow min(ex_{i,j,t} + \Delta ex_{i,j,t}, ul_{ex})$
6. $\forall h_k \in H_c$, do
   i. $\Delta sr_{i,k,t} \leftarrow SRU(CY_{i,k,t,j})$
   ii. If $sr_{i,k,t} + \Delta sr_{i,k,t} < ll_{sr}$ Then
      $sr_{i,k,t} \leftarrow ll_{sr}$
   iii. Else
      $sr_{i,k,t} \leftarrow max(sr_{i,k,t} + \Delta sr_{i,k,t}, ul_{sr})$

4. IMPLEMENTATION

The SimCoL environment was implemented using the Java version of the Repast – a multiagent simulation toolkit. Table 4 summarizes the implementation details of the components of the SimCoL environment. Table 4 describes: (a) the categorizations and the ranges of the randomly generated values in SimCoL, i.e., the task difficulty and the student attributes and (b) the weights and constants used in the equations in Section 3.

Table 4. Categorizations, Distributions, Weights, and Constants Used in SimCoL

	Categorizations and Distributions		
Eq.	Attribute	Categorization	Generated from Normal Distribution with
(1)	Task Difficulty	Low[0.0, 0.3), moderate[0.3, 0.6), and high [0.6,1.0)	$\mu_{di} = 0.5$, $\sigma_{di} = 0.1$, and range [0,1]
(2)	Expertise	Low [0,0.3), moderate [0.3,0.6), and high [0.6,1.0],	$\mu_{ex} = 0.3$, $\sigma_{ex} = 0.25$, and range [0,1]
(2)	Ability	Low [0,0.3), moderate [0.3,0.6), and high [0.6,1.0]	Calculated using (3) with range [0,1]
(2)	Motivation	Low[0,0.2), moderate[0.2,0.8), and high[0.8, 1.0]	$\mu_{mo} = 0.4$, σ_{mo} 0.25, and range [0,1]
(2)	Emotion	Sad[−1.0, −0.5), neutral[−0.5,0.5), and happy[0.5,1.0].	$\mu_{em} = 0$ and , $\sigma_{em} = 0.5$, and range [0,1]
(2)	Social Relationship	Unknown[0, 0.2), familiar[0.2,0.8), and friend[0.8, 1.0]	$\mu_{sr} = 0.4$, $\sigma_{sr} = 0.25$, and range [0,1]
	Weights and Proportionality Constants		

(3)	Weights: $w_{abx} = 0.4$, $w_{abm} = 0.4$, and $w_{abe} = 0.2$
(4)	Proportionality constant: 0.9
(5)	Weights: $w_{omo} = 0.8$ and $w_{gmo} = 0.2$
(6)	Weights: $w_{csr} = 0.5$ and $w_{csq} = 0.5$
(7)	Proportionality constant: 0.001
(8)	Weights: $w_{hab} = 0.8$ and $w_{hde} = 0.2$ and proportionality constant: 0.001
(9)	Proportionality constant: 0.001
(11)	Proportionality constant: 0.1
Other Constants	
Collaboration threshold $\kappa_{ch} = 0.2$	
Zone of proximal development threshold $\kappa_{zone} = 0.99$	
Scaffolding Acceptance Threshold $\kappa_{sc} = 0.2$	

5. ACKNOWLEDGMENTS

The work reported here was partially supported by NSF Grant DMI-044129 and Microsoft Research.

6. REFERENCES

[1] D. Brooks and D. Shell, "Working memory, motivation, and teacher-initiated learning," *Journal of Science Education and Technology*, vol. 15, 2006, pp. 17-30.

[2] K. Issroff and T.D. Soldato, "Incorporating Motivation into Cooperative Learning," *Proc. European Conference on Artificial Intelligence in Education*, 1996, pp. 284-290.

[3] I. Jones and A. Pellegrini, "The Effects of Social Relationship, Writing Media and Microgenetic Development on First-Grade Students\' Written Narratives," *American Educational Research Journal*, vol. 33, 1996, pp. 691-718.

[4] C. Graham and M. Misanchuk, *Computer-mediated learning groups: benefits and challenges to using groupwork in online learning environments: online collaborative learning theory and practice*, Hershey, PA: Information Science, 2004.

[5] P. Dillenbourg, *What do You mean by collaborative learning?*, Oxford: Elsevier, 1999.

[6] S. Puntambekar and R. Hubscher, "Tools for Scaffolding Students in a Complex Learning Environment: What Have We Gained and What Have We Missed?," *Educational Psychologist*, vol. 40, 2005, pp. 1-12.

[7] M.D. E. Sklar and M. Co, "SimEd: Simulating Education through a Multiagent System," *Proc. AAMAS*, 2004, pp. 998-1005.

[8] E. Sklar and M. Davies, "Multiagent simulation of learning environments," *Proc. AAMAS*, 2005, pp. 953-959.

[9] N. Khandaker, L. Soh, and H. Jiang, "Student Learning and Team Formation in a Structured CSCL Environment," *ICCE*, 2006, pp. 185-192.

[10] M. Davies and E. Sklar, "Modeling human learning as a cooperative multiagent interaction," *Proc. AAMAS Workshop on Humans and Multi-Agent Systems*, 2003.

[11] M. Spoelstra and E. Sklar, "Using Simulation to Model and Understand Group Learning," *Proc. AAMAS'07 Workshop on Agent Based Systems for Human Learning and Entertainment*, 2007.

[12] M. Spoelstra and E. Sklar, "Using simulation to model and understand group learning," *International Transactions on Systems Science and Applications*, vol. 4, 2008, pp. 76-88.

[13] D.F. Shell and D.W. Brooks, "The Unified Learning Model: Implications for Learning and Teaching," 2007; http://dwb4.unl.edu/index.html.

[14] L. Soh, N. Khandaker, and H. Jiang, "Multiagent Coalition Formation for Computer-Supported Cooperative Learning," *IAAI*, 2006, pp. 1844-1851.

[15] A. Gogoulou et al., "Forming Homogeneous, Heterogeneous and Mixed Groups of Learners," *ICUM*, 2007, pp. 33-40.

[16] H. Ellis et al., "Emotion, motivation, and text comprehension: the detection of contradictions in passages," *Journal of Experimental Psychology: General*, vol. 126, 1997, pp. 131-146.

[17] A. Inaba et al., "How Can We Form Effective Collaborative Learning Groups?," *Proc. ITS*, 2000, pp. 282-291.

[18] A. Schon, *Educating the Reflective Practitioner*, San Francisco, CA.: Jossey-Bass Publishers, 1987.

[19] A. Collins, J. Brown, and A. Holum, "Cognitive apprenticeship: making thinking visible," *American Educator*, vol. 6, 1991, pp. 38-46.

[20] L. Vygotsky, *Mind in society*, Cambridge, MA: Harvard University Press, 1978.

[21] A. Gokhale, "Collaborative learning enhances critical thinking," *Journal of Education*, vol. 7, 1995, pp. 25-33.

[22] G. Stahl, *Building Collaborative Knowing: Elements of a Social Theory of CSCL, What We Know about CSCL and Implementing It in Higher Education*, Norwell, MA.: Jossey-Bass Publishers, 2004.

[23] K. Kreijns, A. Paul, and W. Jochems, "Identifying the Pitfalls for Social Interaction in Computer-Supported Collaborative Learning Environments: a Review of the Research," *Computers in Human Behavior*, vol. 19, 2003, pp. 335-353.

[24] J.V.D. Linden and P. Renshaw, *Dialogic Learning*, Dordrecht: Kluwer Academic Publishers, 2001.

[25] J. Zumbach, J. Schönemann, and P. Reimann, "Analyzing and supporting collaboration in cooperative computer-mediated communication," *2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!*, Taipei, Taiwan: International Society of the Learning Sciences, 2005, pp. 758-767; http://portal.acm.org/citation.cfm?id=1149293.1149393.

[26] K. Issroff and A. Jones, "Learning Technologies: Affective and Social Issues in Computer-Supported Collaborative Learning," *Computers and Education*, vol. 44, 2005, pp. 395-408.

[27] E. Vass, "Friendship and Collaborative Creative Writing in the Primary Classroom," *Journal of Computer Assisted Learning*, vol. 18, 2002, pp. 102-110.

[28] C. Chalmers and R. Nason, “Group metacognition in a computer-supported collaborative learning environment,” *Proc. International Workshop on Microelectromechanical Systems*, 2005, pp. 35-41.

[29] T. Clear and D. Kassabova, “Motivational patterns in virtual team collaboration,” *7th Australasian conference on Computing education*, 2005, pp. 51-58.

[30] J. Bruner, “Vygotsky: A historical and conceptual perspective.,” *Culture, communication, and cognition: Vygotskian perspectives*, Cambridge, England: Cambridge University Press, 1985, pp. 21-34.

[31] C. Cazden, “Peekaboo as an instructional model: Discourse development at home and at school,” *Stanford Papers and Reports in Child Language Development*, vol. 17, pp. 1-19.

[32] P. Greenfield, “Historical change and cognitive change: A two-decade-follow-up study in Zinacantan, a Maya community in Chiapas, Mexico.,” Biennial Meeting of the Society for Research in Child Development, vol. 6, 1999, pp. 92-98.

[33] A. Harrer, N. Malzahn, and A. Wichmann, “The remote control approach - An architecture for adaptive scripting across collaborative learning environments,” *Journal of Universal Computer Science*, vol. 14, 2008, pp. 148-173.

[34] X. Liu, B. Doore, and L. Li, “Scaffolding Knowledge Co-Construction in Web-based Discussions through Message Labeling,” *Proc. Society for Information Technology and Teacher Education International Conference 2008*, 2008, pp. 3041-3046.

[35] T. Sandholm, “Limitations of the Vickrey auction in computational multiagent systems,” 1996, pp. 299-306.

[36] C. Christodoulopoulos and K. Papapanikolaou, “Investigation of Group Formation Using Low Complexity Algorithms,” *Proc. Personalization in E-Learning Environments at Individual and Group Level Workshop in 11th Intl. Conf. on User Modeling*, 2007, pp. 57-60.

[37] G. Stahl, T. Anderson, and D. Suthers, *Computer-supported collaborative learning: An historical perspective*, 2006; http://www.cis.drexel.edu/faculty/gerry/cscl/CSCL_English.pdf.