1-6-2009

# Multiagent Simulation of Collaboration and Scaffolding of a CSCL 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-2009-0002
Issued Jan. 6, 2009

# Multiagent Simulation of Collaboration and Scaffolding of a CSCL 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

Multiagent techniques improves student learning in Computer-Supported Collaborative Learning (CSCL) environments through multiagent coalition formation and intelligent support to the instructors and students. Researchers designing the multiagent tools and techniques for CSCL environments are often faced with high cost, time, and effort required to investigate the effectiveness of their tools and techniques in large-scale and longitudinal studies in a real-world environment containing human users. Here, we propose SimCoL, a multiagent environment that simulates collaborative learning among students and agents providing support to the teacher and the students. Our goal with SimCoL is to provide a comprehensive testbed for multiagent researchers to investigate (1) theoretical multiagent research issues e.g., coalition formation, multiagent learning, and communication, where humans are involved, and (2) the impact and effectiveness of the design and implementation of various multiagent-based tools and techniques (e.g., multiagent-based human coalition formation) in a real-world, distributed environment containing human users. Our results show that SimCoL (1) closely captures the individual and collective learning behaviors of the students in a CSCL environment, (2) identify the impact of various key elements of the CSCL environment (e.g., student attributes, group formation algorithm) on the collaborative learning of students, (3) compare and contrast the impact of agent-based vs. non-agent-based group formation algorithms, and (4) provide insights into the effectiveness of agent-based instructor support for the students in a CSCL environment.

## Categories and Subject Descriptors

I.6.3 [**Simulation and Modeling];** Model Development; I.6.5 [**Simulation and Modeling]** Applications; I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems.

## General Terms

Algorithms, Design, Experimentation, Human Factors.

## Keywords

Multiagent simulation, Computer-Supported Collaborative Learning.

Multiagent Simulation of Collaboration and Scaffolding of a CSCL Environment, Author(s), *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra, and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. XXX-XXX. 

## 1. INTRODUCTION

Computer-supported collaborative learning (CSCL) environments implement student learning by enhancing their collaborative learning using computer and Internet technologies. Today, CSCL environments contain agents and agent-based services to improve the collaborative learning of students from two different aspects. First, the agents act as assistants to the students by monitoring the difficulties they face and helping them with customized support. Second, the agents act as assistants to the teacher providing decision support and helping him or her with tasks like group formation. To design agents, agent-based services, and agent-based algorithms for a CSCL environment, it is essential to: (1) understand how those various elements of the CSCL environment work together to produce the learning outcome of the students and (2) investigate how those services impact the students' interactions and learning outcomes. Furthermore, without testing their algorithm on a large group of students for a sufficiently long time, it is difficult for the researchers to: (1) fully understand the impact of their designs and (2) evaluate their designs and algorithms against the state of the art. Albeit considered the most authentic way of validating the results, it is often difficult to conduct experiments with human users for various reasons: (1) it is difficult to acquire enough students for long enough time to do the experiments, (2) replication of experiments is often not possible, and (3) experiments may yield unwanted consequences (e.g., student apathy toward the use of CSCL environment) if the agents or agent-based services do not work as expected. One way to alleviate these difficulties is by *agent-based simulation* of the CSCL environment.

However, today's simulation effort of the CSCL environment has yet to consider the role of agents in supporting (or scaffolding) the activities. When designed based on the individual and collaborative learning theories, the students and their interactions with each other in the simulation would closely represent the collaborative learning in the real-world CSCL classroom. Existing tools such as [5] only simulate the student behavior using agents and do *not* include agents that act as the assistant agents or any agent-based services or algorithms. As a result, the decision making process of the CSCL module that provides scaffolding to help both the teacher and the students, as well as the appropriateness and costs of such a module, have not been studied as comprehensively as necessary.

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. In addition, SimCoL environment provides opportunities for multiagent researchers to investigate: (1) theoretical multiagent research issues in *coalition formation*, *multiagent learning*, and *communication* and (2) the design and implementation of various multiagent-based tools and techniques (e.g., the effectiveness of a multiagent-based human coalition formation algorithm [22]) designed for *real-world, distributed* environments containing *human users.*

# 2. LEARNING

In this section, we briefly describe three groups of learning theories we use for our framework (1) *individual learning*, (2) *peer-based learning*, and (3) *collaborative learning*. Here peer-based learning is differentiated from collaborative learning as it focuses on learning involving two learners only. For details on these learning theories, please refer to our technical report [24]. The objective of this section is to identify key observations derived from these leaning theories. These observations drive the design of our simulation, later described in Section 3.

## 2.1 Individual Learning

According to learning theories [1,7] 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 [1], and (4) emotional state [7]. From the above, we draw the following:

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.*

The combined effect of these four components on the learning of a student described by [1] further lead to the following:

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 [8] such as: learning by *observation*, learning by *teaching/guiding*, learning by *being taught*, learning by *reflection/self-expression*, learning by *apprenticeship*, learning by *practice*, and learning by *discussion*. From these peer-based learning scenarios, 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. These observations are summarized in Table 1.

**Table 1. Possible Learning Scenarios among Peers**

| Observation | Student's— Peer's Knowledge | Learning by |
|---|---|---|
| 7 | High—High | Observation, Reflection, Practice and Discussion |
| 8 | High—Low or Low—High | Observation, Teaching, Being Taught Reflection, Practice, and Discussion |
| 9 | 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 [9]. For example, 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 [9]. This gives us our next observation:

Observation 10: *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 [10]. Derived from Stahl [11] are:

Observation 11: *The collaborative knowledge building is a cyclic process that feeds on itself and converges exponentially faster.*

Observation 12: *This collaborative knowledge building cycle is a hermeneutic cycle, meaning, "one can only interpret what one already has an interpretation of".*

Observation 13: *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.*

Kreijns [12] 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 [27].

Observation 14: *The collaboration among the members of a group of students occurs due to their interaction/discourse with each other.*

Zumbach [13] 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 [13] is: (a) student a proposes a solution for the assigned task, (b) student b accepts or proposes another solution to the task. Thus:

Observation 15: *The compilation of discourse/interaction patterns presented by Zumbach et al.* [13] *describes a typical dyadic (between two students) learning scenario in terms of a chain of action-reaction patterns.*

The quality the discourse/interactions within a group depends on the affective state of a student [2] and his or her social relationship with other students in the group. Jones and Issroff [14] and Vass [15] report that, students who are friends have established ways of working which are implicitly understood rather than explicitly discussed. In addition, [12] mentions that social relationships contribute to common understanding, an orientation towards cooperation, and the desire to remain as a group. Finally, as reported in [3], the students form their view of other students due to the type and extent of collaboration they receive from their peers. Clear and Kassabova [16] further report that in collaborative learning settings 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, and vice versa. We derive from the above the following observations:

Observation 16: *Good social relationship and/or friendship improve the quantity and quality of interaction among a group of students.*

Observation 17: *The quantity and quality (i.e., learning outcome) of interactions among a group of students vary 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 18: *Motivation of the group members' impacts the motivation of a student. If the group members are motivated, it may increase that student's motivation, and vice versa.*

Observation 19: *Social relationship between a student and his or her peer (as perceived by the student) change according to the frequency, extent, and quality of collaboration (e.g., how many times did my peer helped me).*

## 2.4 Scaffolding

Bruner [17] and Cazden [18] 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. Over time, the concept of scaffolding has been introduced into traditional classrooms to aid learners to achieve difficult learning objectives and complete difficult tasks [3] where tools and software 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. Researchers in the CSCL community are now utilizing scaffolding in the form of incorporating structure of learning activities (e.g.,[19]) and improving peer support (e.g.[20]). As CSCL researchers (e.g., [3,20] 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. Hence:

Observation 20: *Scaffolding in the CSCL environment can be provided by: (1) providing structure and support for completing tasks and (2) improving of peer support.*

Observation 21: *Scaffolding in the CSCL environment may be used to improve the knowledge of the learners regarding the assigned task.*

Observation 22: *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, $H = \{h_1, \dots, h_{ns}\}$ is a set of agents who represent the students in a collaborative classroom environment, $SA = \{sa_1, \dots, sa_n\}$ 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.

## 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, $T_j = \langle ct_j, di_j, tl_j, \overrightarrow{sq_j} \rangle$ (1). Here, $ct_j$ denotes the concept of the task. This concept represents the subjective knowledge required to solve the task. $di_j \in \mathbb{R}$, is the difficulty of the task as determined by the teacher. $tl_j$ is the time limit within which the task is to be completed. $\overrightarrow{sq_j}$ where is a vector representing the student groups' (who are working on the task) view of the solution quality of the assigned task $T_j$ at time $t$.

## 3.2 Student

We represent the model $hm_{i,t}$ of each student $h_i \in H$ in SimCoL by a 6-tuple: $hm_{i,t} = \langle K_{i,t}, AB_{i,t}, MO_{i,t}, EM_{i,t}, SR_{i,t}, TSQ_{i,t} \rangle$ (2), where, $K_{i,t} = \{(ct_j, ex_{i,j,t}) | \forall T_j \in T\}$ is the knowledge of student $h_i$ at time t with $ct_j$ representing the concept of $T_j$ and $ex_{i,j,t} \in \mathbb{R}$ is 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. $AB_{i,t} \in \mathbb{R}$, is the ability of $h_i$ at time $t$ for task $T_j$. $MO_{i,t} \in \mathbb{R}$, is the motivation of $h_i$ at time $t$. $EM_{i,t} \in \mathbb{R}$, 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 $sr_{i,k,t} \in \mathbb{R}$ is the social relationship between $h_i$ and $h_k$ at time $t$ as perceived by $h_i$. $TSQ_{i,j} \in \mathbb{R}$ denotes the target solution quality of the task $T_j$ of $h_i$ at time $t$.

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 16 and 17. Also, 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: $AB_{i,j,t} = AB(hm_{i,t}) \propto w_{abx} \cdot ex_{i,j,t} + +w_{abm} \cdot MO_{i,t} - w_{abe} \cdot |EM_{i,t}|$ (3), where $w_{abx}, w_{abm}$, and $w_{abe}$ are weights. According to Eq. 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. We also define the target solution quality of a student with: $TSQ_{i,j,t} \propto AB_{i,t}$ (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.

## 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** [21], and **VALCAM** [25] group formation method. Table 2 shows how the teacher carries out the CSCL session through a set of simulation steps. First, the teacher initializes the classroom (tasks, group formation scheme, how often scaffolding should be provided, and how many groups would receive scaffolding). Then, for each initialized task, the teacher: (1) initializes a collaborative session (Step 2a), forms student groups (Step 2b-d), and announces the start of the collaborative session to all students (Step 2e). Then until the collaborative session is over, the teacher periodically sorts the groups according to their current achieved solution quality of the task (Step 2fa(1)) and then selects the groups who have the lowest solution quality. Those selected groups are then provided scaffolding (Step 2fa(2)). Finally, the teacher announces the end of the collaborative session when the time limit for the current task is over (Step 2g).

**Table 2. Simulation Steps of Teacher**

| Simulation Steps of Teacher I |
|---|
| 1. Initialization: $T \leftarrow \{T_1, \dots, T_{nt}\}$, $G_{fs} \leftarrow$group formation scheme, $t_{sc} \leftarrow$scaffolding period, $n_{sc} \leftarrow \lvert G_{sc} \rvert \cdot r_{sc}$, students $H = \{h_1, \dots, h_n\}$, and agents $Sa = \{Sa_1, \dots, Sa_n\}$ |
| 2. For all tasks $T_j \in T$, do, |
| a. Initialize collaborative Session $s_j$: $t \leftarrow 0$, $G \leftarrow G_1, \dots, G_m$, $n_g \leftarrow \lfloor H_m / m \rfloor$, and Announce task $T_j$ to students $H$, |
| b. If $G_{fs} = R_n$, form Random Group for $H$ |
| c. Else If $G_{fs} = H_a$, form Hete-A [21] groups for H |
| d. Else If $G_{fs} = H_v$, form VALCAM [22] groups for H |
| e. Announce start of collaborative Session $s_j$ to $H$ |
| f. While ($t < tl_j$ true) |
| 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. $t \leftarrow t + 1$ |
| g. 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 [22]. In VALCAM, the teacher agent hosts iterative auctions and the student agents bid in those auctions to form student groups that have high average of student expertise and social relationship.

## 3.5 Collaboration and Scaffolding

Following Observations 14 and 15, 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 functions that dictate the behavior of the student agents simulating the collaborative learning in SimCoL. In the following, we assume that two students $h_i$ and $h_k$ with models $hm_{i,t}$ and $hm_{k,t}$ are working in a group $G_m$ to solve task $T_j$

***Motivation Update*** (based on Observation 18): $MSU(MO_{i,t}, G_m) = \left[w_{omo} MO_{i,t} + w_{gmo} \sum_{k \in G_m - h_i} \frac{MO_{k,t}}{(|G_m| - 1)}\right]$ (5) where $w_{omo}$ and $w_{gmo}$ are weights, $MO_{i,t} \in hm_{i,t}$.

***Collaboration Probability*** (based on Observation 16): $CP(hm_{i,t}, hm_{k,t}, T_{j,t}) = w_{csr} \cdot sr_{i,k,t} + w_{csq} \cdot (sq_{j,t} - TSQ_{i,j,t})$ (6) where $w_{csr}$ and $w_{csq}$ are weights, $sr_{i,k,t} \in SR_{i,t}, TSQ_{i,j,t} \in hm_{i,t}$, and $sq_{j,t} \in T_j$.

***Collaboration Cycle*** (based on Observation 15): $cc_{i,k,t,j} = \{act_{i,k,t,j}, rct_{i,k,t,j}, lct_{i,k,t,j}\} \subseteq CC_{i,k,t,j}$ denotes a collaboration cycle completed by $h_i$ with $h_k$ 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 $rct_{i,k,t,j}$. $ci_{i,k,t,j} \subseteq CI_{i,k,t,j}$ denotes a collaboration cycle initiated by $h_i$ but declined by $h_k$. $CY_{i,k,t,j} = \{CC_{i,k,t,j}, \ldots, CC_{k,i,t,j}, \ldots, CI_{k,i,t,j}\}$ denotes the set of all collaboration cycles between $h_i$ and $h_k$ regarding $T_j$.

***Solution Quality Update:*** $SQU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j}) = 0$ If $p_{sq} \geq \kappa_{sq}$ and $\propto (ab_{i,j,t} + ab_{k,j,t})/di_j$ *Otherwise* (7) where $cc_{i,k,t,j}$ is a collaboration cycle, $\kappa_{sq}$, $p_{sq} \in \mathbb{R}$ denotes the solution quality update probability threshold and a random number that is drawn from a uniform random distribution respectively. $ab_{i,j,t} \in AB_{i,t} \in hm_{i,t}$, $ab_{k,j,t} \in AB_{k,t} \in hm_{k,t}$, and $d_j \in T_j$.

***Human Expertise Update*** (based on Observation 3,4,7-10 and 13): $HEU(hm_{i,t}, hm_{k,t}, cc_{i,k,t,j}) = 0\ If\ DE(hm_i, hm_k, T_j) > \kappa_{zone}$ otherwise $\propto [w_{hab}\, ab_{i,j,t} + w_{hde}\ DE(hm_{i,t}, hm_{k,t}, T_j)]$ (8) with $DE(hm_{i,t}, hm_{k,t}, T_j) = |ex_{i,j,t} - ex_{k,j,t}|$ (9) $cc_{i,k,t,j}$ is a collaboration cycle, $\kappa_{zone}$ is the zone of proximal development constant, $w_{hab}$ and $w_{hde}$ are weights, $ab_{i,j,t} \in AB_{i,t}, ex_{i,j,t} \in K_{i,t} \in hm_{i,t}$, $ex_{i,k,t} \in K_{k,t} \in hm_{k,t}$.

***Social Relationship Update*** (based on Observation 19): $SRU(CY_{i,k,t,j}) \propto [|CC_{i,k,t,j}| - |CI_{i,k,t,j}|]/[|CC_{i,k,t,j}| + |CI_{i,k,t,j}|]$ (10) where collaboration cycle sets $CC_{i,k,t,j}, CI_{i,k,t,j} \in CY_{i,k,t,j}$.

***Scaffolding Effect*** (based on Observation 20-22): $SEU(hm_{i,t}, ct_j, SO_j) \propto 1/(1 + |ex_{i,j,t} - sol_{i,j}|)$ If $p_{sca} \geq \kappa_{sc}$ and *0* otherwise (11) where $SO_{i,j} = \langle ct_j, sol_j, csc_j \rangle$ is the scaffolding object, $ct_j \in T_j$, $sol_j$ denotes the level of expertise for the student the scaffolding is designed for, $csc_j$ denotes the cost (e.g., time and effort required to design the object) of the scaffolding, $p_{sca}$ is a probability value drawn from a uniform distribution, and $\kappa_{sc}$ is the scaffolding threshold.

Table 3 shows the simulation steps of a student in SimCoL with the various formulas that are used by the agents in parenthesis. During initialization, the student receives its group assignment and the task (Step 1) from the teacher (Step 2a in Table 2). Then the student updates its own motivation according to other group member's motivations, and its ability. During the session, the student tries to collaborate with its group members if the quality of the solution is less than its expected solution quality (Step 2a) or if someone else in the group wants to collaborate (Step 2b). In both of these cases, whether the collaboration is successful or not depends on the collaboration probability (Step 2b(i)). During the collaborative session, if the student receives scaffolding from the teacher (Step 2c) in the form of a scaffolding object, it updates its expertise. Finally, when the collaborative session ends, the student updates its own view of its social relationship with all its group members (Step 3).

**Table 3. Simulation Steps of Student**

| Simulation Steps of Student $h_i$ |
|---|
| 1. Initialize: group $G_m$, task $T_j$, update motivation (5) and ability (3) |
| 2. Until collaborative Session $s_j$ is over, do, |
| a. 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 |
| Complete and store collaboration cycle in $CC_{i,k,t,j}$ and update solution quality (7), and expertise (8) |
| iii. Else |
| store failed collaboration cycle in $CI_{i,k,t,j}$ |
| b. If received collaboration request from $h_k$ Then |
| i. If $CP(hm_{i,t}, hm_{k,t}, T_{j,t}) > \kappa_{ch}$ Then |
| Complete and store collaboration cycle in $CC_{k,i,t,j}$, and update expertise (8) |
| ii. Else |
| Decline collaboration request from $h_k$ and store failed collaboration cycle in $CI_{k,i,t,j}$ |
| c. If received scaffolding $SO_j$, then |
| Update expertise (11) |
| 3. Update social relationship (10) for group members |

# 4. RESULTS

The SimCoL environment was implemented using the Java version of the multiagent simulation toolkit Repast [26]. The students, the teacher, and the agents are designed as Java objects. The student models in SimCoL were generated randomly from normal distributions with attribute values in the range [0,1] and those values were divided into three equal intervals *low, medium, and high*. The probability distributions and the parameters that govern student and teacher behavior can be accessed through the Repast GUI for running simulation experiments. In this section, we present four aspects of our simulation results: (1) validity, (2) dual attribute analysis to investigate relationship between any pair of attributes, (3) analysis of coalition formation schemes on student learning gains, and (4) scaffolding analysis to investigate its role in supporting collaborative learning.

## 4.1 Validity Analysis

To validate SimCoL, we compared our simulation results with previously published collaborative learning patterns. First, matching [7], we observed that the high-ability students are able to learn at a faster rate (0.33 vs. 0.1) than the low-ability students. Since a student's expertise gain due to collaboration depends upon its own ability, a high-ability student can generally improve the expertise more than a low-ability student, as prescribed in (8). Furthermore, we observed that: (1) the total expertise gain of the students converges to a final value and (2) the rate of change of expertise is higher in the beginning and slows down at the end, coinciding with published reports [8,15]. These provide basic validation for SimCoL. Readers are referred to [32] for the details of our validity analysis.

## 4.2 Dual Attribute Analysis

Dual attribute analysis allows us to: (a) investigate how the students belonging to the different categories of an attribute respond to the changes in another attribute, e.g., how do the student with low expertise react to a change in their motivation, and (b) investigate whether a student's lower value in an attribute can be compensated by a higher value. To collect data for this experiment, we ran the simulation with 10 different simulation runs (with unique seeds) for 100 students for 2000 simulation ticks for each run by varying the values of two attributes at a time and plotted the expertise gain of the students against their changing attribute values. For each plot, we divide the collected data points into three categories of an attribute: low, medium, and high, and then plot the average expertise gain against the remaining attribute for each of those three categories. Due to space consideration, here we report on a subset of the analysis.

Fig. 1 shows the average expertise gain of the students when the average initial expertise and the average motivation of those students are varied. According to Fig 1, we see when the average motivation of the students is increased, the students of all categories (low, medium, and high) of expertise are able to improve their expertise gain. This is to be expected as dictated by Eq. 8 derived in Section 3.5, which in turn is determined by the motivation (Eq. 3). However, with the simulation, we are able to also observe the compound effects of these two factors, as manifested in the rates of changes of expertise gain: 0.17, 0.13, 0.08, respectively, as shown in Fig. 1.

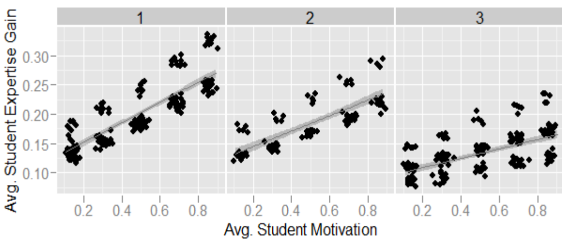


**Fig. 1. Student expertise gain vs. average student motivation for low, medium, and high expertise (left to right) students.**

Fig. 2 shows that as the social relationship of students improves, their expertise gain improves at first, and then that rate of improvement slows down to zero. This occurs due to our use of student social relationship while calculating the collaboration probability among two students (Eq. 6). The expertise gain of the students in the group depends on how well they collaborate. As the social relationship among the students starts to increase from initial lower value, the probability of them collaborating increases. As a result, they are able to gain more expertise. However, when their social relationship values are near maximum and all students in every group are collaborating, increasing the social relationship value further, does not impact their average expertise gain. Once again, Fig. 2 also shows that, the rate at which the expertise gain of the students increased due to the improvement in their average social relationship is slower for students with higher expertise (rate of expertise gain increase: 0.25 (low), 0.17 (medium), and 0.07 (high)). This shows the compound effect of student expertise and student social relationship on the expertise gain of the students. According to Eq. 6, increased social relationship increases the probability that all group members collaborate with each other. However, increasing the average expertise decreases the expertise difference among the students which in turn reduces their overall expertise gain (Eq. 8). As a result, increasing the social relationship produces less impact when the average expertise of the students is high.

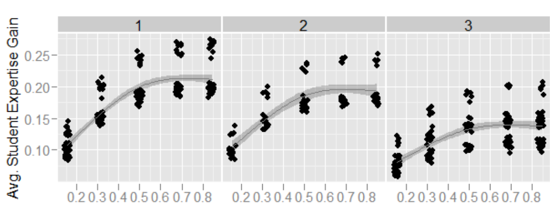


**Fig. 2. Student expertise gain vs. average student social relationship for low, medium and high expertise (left to right) students**

In conclusion, the important observation regarding our dual attribute analysis is that there are components in the collaborative learning environment that impact one another's effect on the expertise gain of the students. This indicates the following: first, for an instructor, it is important to understand how these components impact one another to achieve optimal expertise gain for the students. Second, a simulation environment like SimCoL may help the instructor gain a better understanding of the *compound* effects of students attributes leading to a more effective design of the CSCL environment.

## 4.3 Analysis of Coalition Formation Schemes

In this section, we study the effect of two teacher-controlled aspects of a typical CSCL environment, i.e., (1) the group size and (2) the group formation scheme, on the average expertise gained by the students. We ran the simulation with expertise distribution mean $\mu_{ex} = 0.8$, expertise distribution standard deviation $\sigma_{ex} = 0.8$, and collaboration threshold $\kappa_{ch} = 0.5$, for a set of 180 students, for 10 different tasks and for 2000 simulation ticks. We also set other parameters to default values representing a typical classroom [25]. During the simulation, the student groups in this experiment were formed using **Random**, **Hete-A**, and **VALCAM** group formation methods with the group size selected from the range of [2,4]. VALCAM is an agent-based algorithm of group formation in which the individual agents bid in an iterative auction to form student groups. While bidding, the agents try to join a group that contains students with: high-expertise and high social relationship values. Hete-A algorithm is a non-agent-based algorithm that forms heterogeneous groups. In Hete-A, the students are first categorized by assigning them to a matrix whose dimensions represent the attributes of a student. Once the students are categorized, the Hete-A algorithm builds heterogeneous groups by selecting students with the highest difference of attribute values according to their position in the matrix. In our implementation, the Hete-A algorithm was used to form groups with the motivation and expertise of the students as the two dimensions of the matrix. Again, we replicated each experiment for 10 different random simulation seeds.

Fig. 4 shows that the students in the group formation algorithm achieved similar expertise gains. However, relatively higher expertise gain of the VALCAM groups against randomly formed groups has been reported by in [22]. This difference in our simulation observation and the reported empirical results can be explained from the viewpoint of our collaboration modeling of the students. In the reported study [22], students in the VALCAM groups changed their group membership a lot less frequently than the students in the randomly-formed groups. As a result, during the limited number of collaborative sessions, the VALCAM-group students were able to: (1) get more familiarized with each other and (2) learn to coordinate their actions better than the randomly formed group's students over time. As a result, in the classroom, the students in VALCAM groups were able to gain higher expertise in the reported experiments. However, while modeling the collaboration of the students in SimCoL, we do not account for this *accumulated familiarity effect* of the students. As a result, VALCAM did not perform better than the other two group formation algorithms. This experiment hints the following: first, although the CSCL literature we have reviewed do not explicitly mention this *accumulated familiarity effect* among the students in CSCL groups, this effect could actually differentiate the expertise gains of students in groups formed by different group formation algorithms. Second, these results indicate that both CSCL and multiagent research community can use SimCoL to: (1) compare and contrast the effectiveness of various agent-based and non-agent-based group formation algorithms through replicable simulations and (2) validate the basic CSCL theories against the reported empirical results and gain valuable insights into both areas.

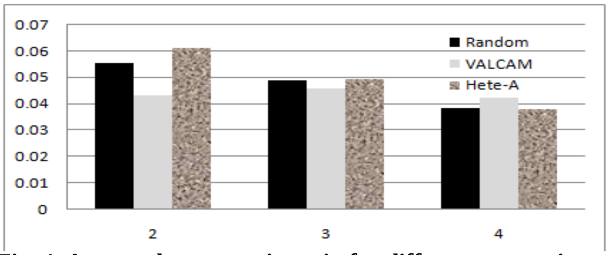


**Fig. 4. Avg. student expertise gain for different group sizes.**

## 4.4 Scaffolding Analysis

In this experiment, we investigate how the individual and group scaffolding improves the expertise of the students when they are collaborating in various types of groups. To collect data for this experiment, we ran the simulation with the same default set of parameters used in Section 4.3 with 10 different simulation seeds for 180 students for 2000 simulation ticks. We calculated: (1) the average improvement in the expertise gain of the students and (2) the cost incurred for providing scaffolding for individuals and groups. For a group in this experiment, one scaffolding object is used *per group* for group scaffolding (i.e., scaffolding cost is required for one scaffolding object) and one scaffolding object *per group member* (i.e., scaffolding cost is equal to the sum of all generated scaffolding objects) is used for individual scaffolding. Fig. 5(a) shows the average improvement of student expertise gains of the students when they are working in random, Hete-A, and VALCAM formed groups. Fig. 5(a) shows that the students in all groups are able to improve their expertise more from the individual scaffolding than from the group scaffolding. This is expected, since: (1) individual scaffolding is designed to address an individual student's needs, and (2) according to our design of scaffolding (Eq. 11), a student's expertise is improved most when the scaffolding is targeted towards his or her expertise level.

Fig. 5(b) shows that for all three types of groups, the group scaffolding yielded more expertise gain per unit cost than the individual scaffolding. The cost of scaffolding denotes the time and effort required for providing scaffolding to the students. Providing individual scaffolding requires more cost since each individual student has to be modeled and different types of scaffolding have to be provided to the students according to their expertise level. On the other hand, group scaffolding requires less cost since the scaffolding action is more generic and only one type of scaffolding is provided to the entire group. But unexpectedly, *the group scaffolding is shown to be more economical in terms of expertise improvement per unit cost*. Upon closer analysis, this can be explained by the cyclic and convergent nature of the collaborative knowledge building process (Observation 11). Due to this cyclic nature, collaborative knowledge is transferred among the group members due to their interactions throughout the collaborative session. Furthermore, our non-adaptive scaffolding process periodically provides scaffolding to a fixed number of student groups by first sorting them according to their performances. However, near the end of the collaborative cycle, due to the heterogeneous nature of groups of the random, Hete-A, and VALCAM groups, there are some students who have already reached near-maximum expertise level. So, scaffolding for such group members is no longer effective. As a result, both individual and group scaffolding do not yield any expertise improvement for those high-expertise group members. But, for those high-expertise group members, the individual scaffolding incurs a much higher cost than would the group scaffolding . As a result, the improvement of expertise *per unit cost* for individual scaffolding is smaller than the group scaffolding. These results show us the improvement-cost tradeoff that occurs for individual and group scaffolding. Although targeted individual scaffolding may improve the expertise gain of a set of students more than group-based scaffolding, the former is less-economical when applied in a non-adaptive manner. With SimCoL, one would be able to pinpoint with higher precision when group and individual learning would be cost-effective for his or her classroom.

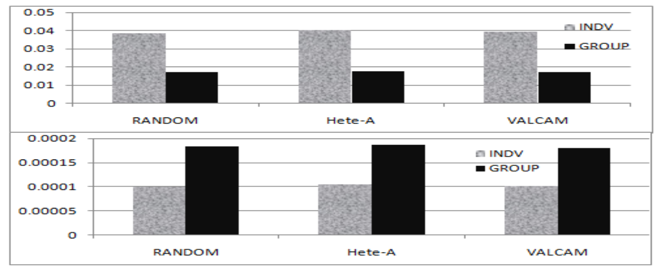


**Fig. 5. (a) (top) Average expertise gain for individual and group scaffolding (b) (bottom) Average expertise gain per unit cost for individual and group scaffolding.**

# 5. RELATED WORK

Sklar and Davies [4] described a simulation environment for the education system called SimEd where they mainly focused on learning from the teacher instead of learning from the peers common in CSCL environments. Spoelstra and Sklar [5] used multiagent simulation to model individuals participating in various group learning scenarios. The researchers used ability, motivation, existing knowledge, and likeliness to help model individual students. Although, the researchers studied the effect of group reward and group composition (i.e., heterogeneity of the members) on the learning outcome of groups, they did not investigate the effects of group formation method, and agent support on the learners. In addition, the researchers did not consider the effect of a student's own ability on his or her learning outcome. However, as CSCL researchers have suggested, (1) the group formation method [23] and agent support [22] can have a significant impact on the collaborative learning outcome of students and (2) so can a student's ability [6] on the individual learning outcome of students. Therefore, we consider these important factors to build a simulation environment that better depicts the collaborative learning of a group of students.

# 6. CONCLUSIONS

Both the learning theories and the techniques used to realize CSCL systems are evolving [23]. This evolving domain implies that it would be useful for the researchers and teachers to have a tool to test those evolving theories and techniques. As a low-cost alternative, simulation-based environments could be used to validate or investigate the usefulness of the CSCL techniques, or in the least, provide hints and guidance to instructors or education researchers on student pedagogy and instructional approaches. In this paper, we have proposed SimCoL, an agent-based tool for simulating the learning process in a CSCL system. We have described the design and implementation of the SimCoL environment and its agents using observations reported by the researchers working in the individual, peer-based and collaborative learning domains. The overall simulation results of the SimCoL environment is consistent with previously reported collaborative learning patterns. Furthermore, our results hint that the SimCoL environment allow the researchers to gain better insights into the impact of: (1) *individual* student attributes, (2) various agent-based and non-agent based group formation algorithms, (3) different types of scaffolding processes on the collaborative learning outcome of students, and (4) CSCL and collaborative learning on real classrooms in particular, and any human-computer environments where online collaborative activities take place among users with diverse behaviors

Our future work involves improving the SimCoL environment by: (1) implementing and analyzing the impact and the cost of providing agent-based collaborative support to students in a CSCL classroom, and (3) improving our modeling of the collaboration of students by accommodating the accumulated familiarity effect. We also plan to validate SimCoL further by running additional simulations on reported CSCL studies.

## 7. ACKNOWLEDGMENTS

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

## 8. REFERENCES

[1] Brooks, D. and Shell, D. 2006. Working memory, motivation, and teacher-initiated learning. J. Sci. Edu. & Tech., vol. 15, 17-30.

[2] Issrof, K. and Soldato, T.D. 1996. Incorporating Motivation into Cooperative Learning, Proc. European Conf. AI in Education, 284-290.

[3] Khandaker, N. Soh, L-K., and Jiang, H. 2006. Student Learning and Team Formation in a Structured CSCL Environment, Proc. ICCE, 185-192.

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

[5] Spoelstra, M. and Sklar, E. 2008. Using simulation to model and understand group learning, Int. Trans. Systems Science & Apps., vol. 4, 76-88.

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

[7] Ellis, H. et al. 1997. Emotion, motivation, and text comprehension: the detection of contradictions in passages, J. Exp. Pysch.: General, vol. 126, 131-146.

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

[9] Vygotsky, L. 1978. Mind in Society, Cambridge, MA: Harvard University Press.

[10] Dillenbourg. P. 1999. What do You Mean by Collaborative Learning?, Oxford: Elsevier.

[11] Stahl, G. 2004. 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.

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

[13] Zumbach, J., Schönemann, J., and Reimann, P. 2005. Analyzing and supporting collaboration in cooperative computer-mediated communication, 2005 Conf. Computer Support for Collaborative Learning: Learning 2005: the next 10 years!, 758-767.

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

[15] Vass, E. 2002. Friendship and Collaborative Creative Writing in the Primary Classroom, J. Computer Assisted Learning, vol. 18, 102-110.

[16] Clear, T. and Kassabova, D. 2005. Motivational patterns in virtual team collaboration, 7th Australasian Conf. Computing Education, 51-58.

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

[18] Cazden, C. 1983. Peekaboo as an instructional model: Discourse development at home and at school, Stanford Papers and Reports in Child Language Development, vol. 17, 1-19.

[19] A. Harrer, N. Malzahn, and A. Wichmann. 2008. The remote control approach - An architecture for adaptive scripting across collaborative learning environments, J. Universal Comp. Sci., vol. 14, 148-173.

[20] Liu, X., Doore, B., and Li, L. 2008. Scaffolding Knowledge Co-Construction in Web-based Discussions through Message Labeling, Proc. Society for Inf. Tech. & Teacher Education Int. Conf., 3041-3046.

[21] Gogoulou, A. et al. 2007. Forming Homogeneous, Heterogeneous and Mixed Groups of Learners, Proc. ICUM, 2007, 33-40.

[22] Soh, L-K., Khandaker, N., and Jiang, H. 2006. Multiagent Coalition Formation for Computer-Supported Cooperative Learning, Proc. IAAI, 1844-1851.

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

[24] Khandaker, N. and Soh, L-K. 2008. On Incorporating Learning Theories to Simulate a Computer-Supported Collaborative Learning Environment. Technical Report TR-UNL-CSE-2008-0006, Department of Computer Science and Engineering, University of Nebraska-Lincoln.

[25] Khandaker, N. and Soh, L-K. 2008. SimCoL: A Simulation Tool for Computer-Supported Collaborative Learning. Submitted to IEEE Transactions on Learning Technologies, 2008.

[26] Repast. 2008. Recursive Porus Agent Simulation Toolkit. http://repast.sourceforge.net.

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