Improving Group Selection and Assessment in an Asynchronous Collaborative Writing Application

Nobel Khandaker, Leen-Kiat Soh
Department of Computer Science and Engineering, University of Nebraska
256 Avery Hall, Lincoln NE 68588-0115 USA
{knobel,lksoh}@cse.unl.edu

Abstract. Two critical issues of the typical computer-supported collaborative learning (CSCL) systems are inappropriate selection of student groups and inaccurate assessment of individual contributions of the group members. Inappropriate selection of student groups often leads to ineffective and inefficient collaboration, while inaccurate assessment of individual contributions of the group members (1) hinders healthy working relationships among members and (2) prevents teachers from providing precise interventions to specific students. To address these issues, our proposed *i*HUCOFS framework forms student groups by balancing the students' competence (what the students know) and compatibility (whom they like as peers) for each group. The competence and compatibility are calculated using the assessment of student contributions derived from a newly implemented asynchronous collaborative writing module's detailed tracking information. Results suggest that: (1) the use *i*HUCOFS framework may improve: (a) the effectiveness and efficiency of the groups, (b) the perception of the students of their peers and their groups, and (c) the collaboration among students with low and high competence and (2) the teacher can use the detailed information tracked by the collaborative writing module to: (a) improve the design of the CSCL tools and (b) provide precise intervention to improve collaboration among the students.

Keywords. Computer-Supported Collaborative Learning, Group Formation, Multiagent System

Although computer-supported collaborative learning (CSCL) systems can improve collaboration and learning among the students, there are several challenges that may discourage a teacher from deploying a CSCL system in his or her classroom. These challenges include the (1) selection of student groups and (2) accurate assessment of individual contributions of the members within the student groups toward the final output of the group (Roberts & McInnerney 2007). These two interrelated issues are important since assigning a student in a group where he or she cannot collaborate or learn and or unfair assessment of a student's individual contributions usually diminish the learning outcomes in a CSCL environment (Roberts & McInnerney 2007).

However, typical CSCL environments and CSCL group formation frameworks do not adequately address these two challenges. For example, typical CSCL group formation methods (Graf & Bekele (2006); Muhlenbrock (2006); Christodoulopoulos et al. (2007); Wang et al. (2007)) do not consider attributes like *compatibility* among group members which has been proven to be important in recent CSCL research (Chalmers & Nason 2005). In addition, typical CSCL environments often do not have any adaptation or learning components which could *capture* and *utilize* the changing student behavior to form better groups over time. Finally, typical CSCL environments or collaborative working environments (e.g., Israel 2007; Erkens et al. 2005; Gogoulou et al. 2005; Constantio-Gonzalez 2003; Teixeira et al. 2002; Vassileva et al. 2002), or typical group formation methodology in CSCL or collaborative learning systems (e.g., Redmond 2001; Graf and Bekele 2006; Muhlenbrock 2006; Christodoulopoulos et al. 2007; Wang et al. 2007) do not discuss accurate assessment of individual student performances which may also impact the group selection process.

Keeping these limitations of the typical CSCL systems in mind, the goal of our research is twofold, we aim to (1) utilize the tracking and modeling capabilities of multiagent systems to improve the assessment of individual student contributions towards their group's solution of the collaborative task and (2) design an adaptive (i.e., able to adapt to changing student behavior) multiagent group formation framework that is able to combine students' knowledge or competence (measured by the improved student assessment process) and compatibility to form better groups over time. As an initial step toward our goal, we present the improved *i*HUCOFS framework for CSCL group formation and its implementation in I-MINDS – a multiagent CSCL environment. In particular, we discuss how we have improved the individual assessment of students while they collaborated to write essays using I-MINDS' Asynchronous Collaborative Writing (ACW) module. Furthermore, we discuss how that improved assessment was used by *i*HUCOFS framework to form better student groups over time.

Note that previously, we have reported on the I-MINDS structure and pedagogical studies in (Soh et al. 2004; Soh et al. 2005; Soh et al. 2006a; Soh et al. 2006b; Soh et al. 2008). Furthermore, the recent improvements of the I-MINDS framework and the details of the *i*HUCOFS framework (i.e., the VALCAM algorithm) have been published in (Soh et al. 2008). Finally, (Khandaker & Soh 2008) contains our preliminary, theoretical version of *i*HUCOFS and its *partial* implementation in *structured* student collaboration in the form of Jigsaw groups. To improve the *i*HUCOFS framework from its preliminary version described in (Khandaker and Soh 2008): (1) we have introduced new measures (e.g., effectiveness and efficiency of formed groups) in *i*HUCOFS to better analyze its performance in forming student groups, (2) we have designed and implemented the ACW module that allows *i*HUCOFS to better calculate the competence and compatibility of the students, and (3) we have added non-structured collaborative features to *i*HUCOFS through the ACW module. More specifically, our newly added ACW module allows the students to collaboratively complete writing assignments and the *i*HUCOFS framework allows the teacher to form student groups by balancing the *competence* (knowledge of the assigned topic) and *compatibility* (preference of group members) of the students.

Finally, we present and discuss the results from our 12-week-long experiment that was designed to study the effects of the new ACW module and the refined *i*HUCOFS framework on the individual contribution of the students toward their groups' work and the collaborative outcome of student groups (e.g., how well they were able to solve the problem), respectively. In brief, the analysis of the data in our experiment suggests that the *i*HUCOFS framework can: (a) improve the effectiveness and efficiency of the student groups, (b) improve the perceptions of the students of their peers and their groups, and (c) improve collaboration among students with low and high competence. Furthermore, our results suggest that the detailed tracking information extracted by the ACW module may allow the teacher to better understand student behavior leading to: (1) improvement of the design of the CSCL tools (e.g., the ACW module) and (2) *precise* intervention to improve the quality of collaboration, i.e., intervention to reduce free-riding among students in our case.

In the following sections, we first briefly discuss the impact of group formation and individual assessment on the collaborative learning outcome of students and thereby motivate the need for better assessment techniques and better group formation methods. Then we describe our theoretical framework and outline our solution approach toward achieving our goals. After that, we discuss the I-MINDS environment and briefly describe the *i*HUCOFS framework and the ACW module. Then, we discuss the implementation of I-MINDS, *i*HUCOFS framework and ACW module. After implementation, we describe our detailed experiment setup including our aim of the experiment and discuss how and to what extent the use of the *i*HUCOFS framework and the ACW module contributed to the alleviations of the two discussed shortcomings of typical CSCL systems. Finally, we present some related research work and our conclusions.

IMPACT of GROUP FORMATION and INDIVIDUAL ASSESSMENT on COLLABORATIVE LEARNING

Accurate assessment of the students' contributions is a critical component of a CSCL environment. The assessment is required to model the attributes (e.g., competence and compatibility) of students to allow the group formation method to form effective and efficient student groups. Furthermore, the information gathered from the assessment mechanism can be used by the teacher to: (1) motivate the students by providing an *explicit* scoring scheme that grades the students according to their *individual* contributions to their group's effort, (2) improve the collaboration among the participating students by better understanding each student's contribution to his or her group's effort, and (3) refine and or revise the design and implementation of the tools and techniques used in the CSCL environment. When an explicit scoring scheme—that uses the detailed tracked information of the students' activities—is used to grade students' performances in a CSCL environment, the students are more likely to be motivated to collaborate with their peers (Roberts & McInnerney 2007) as they know their contributions can be held accountable. Furthermore, the teacher can use the detailed tracked information to improve the collaboration of the students through *precise* intervention. For example, the teacher can figure out whether any student is free-riding (i.e., not sharing the group's workload but receiving scored because of membership to the group) and intervene accordingly, which is another common problem in typical CSCL environments (Roberts & McInnerney 2007). Finally, by closely inspecting the behavior patterns of the students in the CSCL classroom, the teacher is able to better understand how the students are using the available CSCL features to collaborate. Then the teacher can remove features that the students do not use or refine the design of the existing ones to further improve future collaboration.

The selection process of student groups is important since in CSCL settings, learning occurs through student collaboration and the quality (whether the participating students were able to learn from each other) and quantity (how many collaborative sessions the students have) of their collaboration largely depends on student attributes like the students' knowledge (i.e., their competence) (Teasley; S. & Roschelle 1993 as cited in Soller et al. 1999) and their compatibility (or social relationship) (Chalmers and Nason 2005; Issroff and Jones 2005). This implies, student groups that contain the members who (1) possess the required problem solving skills and (2) are compatible with each other would be collaborate well. As a result, the improved collaboration among the former group's members would lead to a better learning outcome. Furthermore, depending on the environment and task, these student-attribute values change (although at different rates). For example, students acquire new skills, develop new friendships while they are participating in collaborative sessions, and grow out of old friendships. As a result, a CSCL group formation algorithm needs to track and model the changing student behavior and try to form better student groups.

Although accurate individual assessment of individual contribution of students and balance of student competence and compatibility are important, as alluded to in the introduction, neither the typical CSCL environments nor the CSCL group formation methods adequately address these two issues.

THEORITICAL FRAMEWORK

Our solution approach toward implementing our goals is composed of two components. The first component of our solution is the asynchronous collaborative writing (ACW) module in I-MINDS. The ACW module allows the students to complete collaborative writing assignment while capturing their actions in detail to track the individual contributions of the students (e.g., number of editions posted) using the underlying I-MINDS agent. Furthermore, the ACW module uses I-MINDS communication functionalities to track the group members' communications about both task-dependent and task-independent topics. When the writing assignment is completed, the ACW module produces a summary that contains a detail report describing *individual contributions* of the students and their *peer-evaluations* of each other. This summary of contribution helps us achieve three objectives. First, the information allows the teacher to design an *explicit* scoring scheme. This scoring scheme then motivates the students to pay attention to collaborate and contribute more. Second, the tracked information about the individual contributions allows the teacher to better understand student behavior patterns leading to: (1) further improvement of the CSCL tools and techniques by using the insights related to the students' usage of those tools and (2) *precise* intervention to improve the quality of collaboration among the group members. Third, the *teacher's evaluation* (which is derived from the tracked information) and the *peer-evaluations* (which is collected by the ACW module in the form of surveys (Soh 2004)) of the students are used in the *i*HUCOFS framework to more accurately calculate their *competence* and *compatibility* (respectively), leading to formation of better student groups.

The second component of our solution is using the *i*HUCOFS group formation framework in I-MINDS to automatically form student groups. *i*HUCOFS group formation framework uses an adaptive student model that consists of that student's competence and his or her compatibility with others to form student groups that contain competent members who are willing to collaborate with their peers. As a result, *i*HUCOFS is able to form groups that (1) are effective and efficient for the current task and (2) improves student collaboration for the current task and thus improving their task-dependent and task-independent knowledge for the future tasks.

I-MINDS

I-MINDS (Intelligent Multiagent Infrastructure for Distributed Systems in Education) employs a set of intelligent software agents, representing individual students and the teacher (or teaching resource in the case of an asynchronous course or lesson) to realize a CSCL environment. The rationale behind using multiagent intelligence is the agent's persistence in tracking and monitoring its environment (student and teacher activities), autonomy in decision making, and responsiveness in providing services to both students and the teacher. These are properties that are useful for distance learning and large CSCL classrooms.

Briefly, in I-MINDS, each student has a personal assistant agent (a student agent), and each teacher has a personal assistant agent (a teacher agent). All these agents interact with their respective users as well as among themselves. These agents exchange information, coordinate their actions, and track inter-agent activities behind the scene. A detailed description of the I-MINDS agents and their capabilities can be found in our recent publication Soh et al. 2008. Here, a brief overview is provided as follows.

Agents: I-MINDS has two types of reactive (Wooldridge 2000) intelligent agents: (1) teacher agents and (2) student agents. These agents are composed of multiple modules and are designed to assist the teacher, the students, and the student groups to realize their collaborative learning goals in the CSCL environment. Notice that the intelligent behavior of the agents in I-MINDS arises from their autonomous responses to the observed teacher and student behaviors.

1. Teacher Agent: A teacher agent helps the teacher by, distributing information to student agents, maintaining profiles for all students, assessing the progress and participation of different students, ranking and filtering the questions asked by the students and forming student groups using the *i*HUCOFS framework. Fig. 1 shows the conceptual modules of the teacher agent in I-MINDS.
2. Student Agent: A student agent, on the other hand, works as a personal helper to a student and provides services that allow him or her to communicate with other students and with the teacher. The student agent also presents the teacher-supplied learning material to the student and forms groups for the assigned student by communicating with other student agents and the teacher agent. Fig. 2 shows the conceptual modules of I-MINDS student agent.

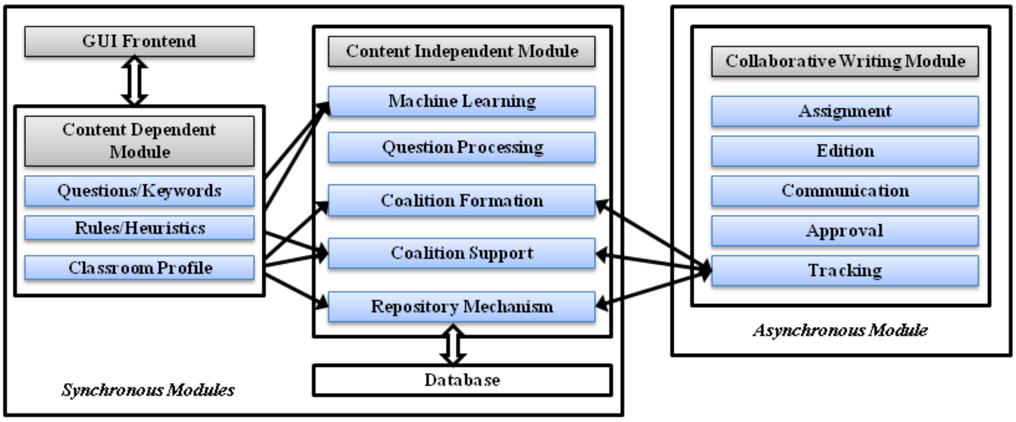

Fig. 1. Modules of an I-MINDS teacher agent. The new asynchronous module has been added to the previously reported synchronous modules (Soh et al. 2008) of I-MINDS.

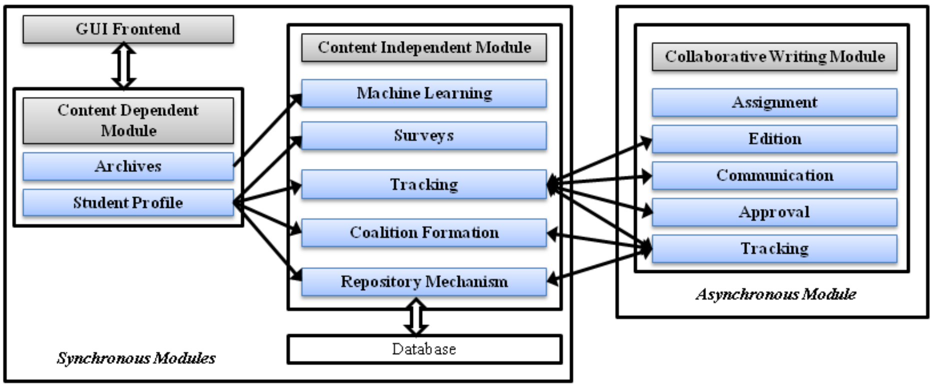

Fig. 2. Modules of an I-MINDS student agent. The new asynchronous module has been added to the previously reported synchronous modules (Soh et al. 2008) of I-MINDS.

Asynchronous Collaborative Writing (ACW) Module

We have designed and implemented the Asynchronous Collaborative Writing (ACW) module as an extension of our current implementation of I-MINDS. While the conceptual versions of the teacher agent and student agent modules of I-MINDS have been discussed in (Soh et al. 2008), here we focus our discussions on ACW module as part of these agent modules. The ACW module uses the built-in agent framework (Figures 2, 3) and repository mechanism of I-MINDS and consists of the following components: (1) assignment, (2) edition, (3) communication, (4) approval, and (5) tracking.

The assignment component in the ACW module allows the teacher to *assign*, *view*, and *archive* collaborative writing assignments to the student groups. While assigning a collaborative writing topic, the teacher can also specify its structure by dividing it into sections. This division of the writing assignment into several sections is designed to achieve the following: (1) it allows the ACW module to track the individual activities or contribution of the students in *greater detail* (i.e., student activities for each smaller section instead of the whole writing assignment can be tracked) and (2) it allows the teacher to help the students to divide the entire writing tasks into several smaller sub-tasks promoting *easier task sharing* among group members. Furthermore, the teacher can specify the constraints related to the writing assignment such as the word limit for the assignment and the assignment due date. The word limit constraint is designed to enable the teacher to encourage collaboration among the students. To complete a collaborative writing assignment that contains all the required sections and that is within the word limit, it is expected that the students are more inclined or compelled to monitor and edit each other's contributions and communicate with each other.

The communication component of the ACW module consists of a chat tool and a forum that are used to track the *task-independent* and *task-dependent* communications among the members of the groups, respectively. The task-independent communications allow the students to discuss matters that are not related to the collaborative writing assignment. Examples of task-independent communication could include: discussions about the topics taught in the classroom, discussions about the usefulness of the collaborative work environment, or even just normal chitchatting among students on personal matters. On the other hand, the task-dependent communications allow the students to post specific comments about the different sections of the collaborative writing assignment. Examples of task-dependent communications could include: comments about the logical flow of a section, discussions regarding task sharing (e.g., who is going to write section 1 and who will edit that later) among the members.

The edition component allows the students to post their contributions to their groups' current version of the collaborative writing assignment. Specifically, once the collaborative writing topic is assigned, the students can contribute in the following ways: (1) propose a new edition for a section (PS), (2) reject the current version of a section (RJ), (3) revise the other students' prepared version of a section (RV), (4) extend the existing version of a section (EX), (5) accept the existing version of a section (AC). With these specific actions, not only do the students have a formal set of actions to facilitate effective collaborations, but also the teacher is able to monitor and realize exactly how each student contributes to the collaborative writing process. To illustrate, not all students think and work the same way. Some students are good at coming up with new ideas (i.e., proposition PS) whereas others are good at revising (RV) or extending (EX) someone else's ideas etc. So, these five different types of editions allow a group of students with different strengths and weaknesses to contribute to the collaborative writing assignment. Subsequently, the teacher is also able to provide specific and precise intervention when a group is lacking in any of the above areas.

The approval component requires each group member to approve the final version of their assignment before it is submitted to the teacher. Furthermore, any edition of the approved version of a writing assignment nullifies the previous approvals and requires every group member to approve again. A typical collaborative writing assignment may involve several editions/modifications by the group members throughout its active period and may lead to several different student views of the most updated version. The approval requirement ensures that all group members are at least held accountable for what they submit as a group and hopefully aware of the final version of the collaborative writing assignment.

While the students are collaboratively writing an assignment using ACW module, their assigned student agent monitors their behavior from four different dimensions: (1) communication, (2) edition, (3) performance, and (4) perception. These four dimensions of student behavior tracking are designed to achieve *two* objectives: (1) to allow *i*HUCOFS' group formation process to track the performance (i.e., competence and compatibility) of individual students and student groups and (2) to enable the teacher more *accurately monitor* individual student contributions and collaboration among group members so that he or she can use precise intervention to improve collaboration (e.g., intervention to discourage free-riding). Table 1 lists the four dimensions of the tracking component and the individual tracked variables for each of those dimensions. Notice that the *communication* and the *edition* dimensions consist of tracked data related to student behavior while they are collaborating, the *performance* dimension consists of the teacher's evaluations of the students and groups (based on the data collected from the communication and edition dimensions) and the *perception* dimension consists of data collected by administering surveys to students.

Table 1: Four dimensions of tracking in the ACW module

Dimension	**Description**	**Tracked Variables**	**Variables' Contribution**
Communication	Student communication through the chat and forum messages	• *chatMsgCount* – Number of chat messages posted by the student • *forumMsgCount* – Number of forum messages posted by the student	• *chatMsgCount* is used by the teacher as an estimate of how well the members of a group are collaborating. • *forumMsgCount* is used to calculate the individual effort of students in *i*HUCOFS' group formation algorithm (Step U4(ii) in Figure 4).
Edition*	Student activities when he or she was modifying an assigned topic summary	• *propositionCount* – Number of propositions posted by a student for the current topic summary • *acceptCount* – Number of accepts posted by a student for the current topic summary • *reviseCount* – Number of revisions posted by a student for the current topic summary • *rejectCount* – Number of rejections posted by a student for the current topic summary • *extensionCount*** – Number of extensions posted by a student for the current topic summary	All of the tracked variables are used to calculate the individual effort of students in *i*HUCOFS' group formation algorithm. Furthermore, these tracked variables allow the teacher to estimate a student's contribution to his or her group's final version of the collaborative writing assignment.

Perfor-mance	The teacher's evaluation score of a user based on his or her individual contribution and the group's topic summary	• *groupEvaluationScore* – Evaluation score of the group for the current topic summary • *individualEvaluationScore* – Evaluation score of the student for the current topic summary	• *groupEvaluationScore* is used as the group reward in *i*HUCOFS' group formation algorithm. • *evaluationScore* is used as the individual student reward in *i*HUCOFS' group formation algorithm (Step S5(iii) in Figure 4).
Perception	How a user evaluates (e.g., through a survey (Soh 2004)) his or her peers and group	• *peerRating* – A student's evaluation of his or her peer's contribution to the topic summary • *teamEfficacyRating* – A student's evaluation of how well his or her group members worked together as a group	• *peerRating* is used to calculate the compatibility between two students in *i*HUCOFS' group formation algorithm. • *peerRating* and *teamEfficacyRating* also allows the teacher to estimate how well they are collaborating, e.g., find answers to questions such as "Are there members in a group who are not contributing?"

*Notice that here each edition count (e.g. proposition count) represents by one single unit of text submission by a student. This text could be of any length.
**Extension action differs from the revision since in extension, students are not able to edit the existing text, only add text to the existing text.

*i*HUCOFS Framework

*i*HUCOFS is a multiagent framework (introduced in Soh and Khandaker 2007 and described in detail in Khandaker and Soh 2008) designed to form and support collaborative learning groups in a CSCL environment that encourages collaboration and improved knowledge gain of students. Researchers suggest that CSCL environment usually provides learning opportunities for the students through *collaboration* with their peers (e.g., learning by teaching, learning by observing, (Inaba et al. 2000)). When the collaborating members possess the necessary knowledge or skill, such collaborations improve the knowledge of the participants and help them learn how to solve the assigned problem (Teasley, S. & Roschelle 1993 as cited in Soller et al. 1999). Furthermore, the quality and quantity of collaboration is often impacted by the peer relationship of the participating students (Chalmers and Nason 2005; Issroff and Jones 2005). The central idea behind *i*HUCOFS is to use the tracking, modeling, autonomous and distributed reasoning capabilities of multiagent systems to track and model the student attributes to form and support (although we do not elaborate on the support aspect in this paper) student groups better. In brief, *i*HUCOFS framework consists of a set of intelligent agents that track and model the students and help the teacher and the students form student groups that contains competent users who are also compatible. Including competent students in a student group allows the not-so-competent students to learn from them peers and having compatible students encourages more collaboration among them. Once the groups are formed, these agents monitor the collaborative activities of the students and also periodically gather information through direct interaction (e.g., surveys). When the collaboration is over, *i*HUCOFS agents utilize the teacher's evaluation of the students (calculated from the monitored information) to form student groups for the next round of CSCL activities. Notice that due to continuous monitoring of collaborative activities, continual interaction with students, and consideration of teacher's assessment of student performance, *i*HUCOFS aims to capture the change in student knowledge and compatibility to form better groups over time. Next, we describe *i*HUCOFS agents' design and architecture, and elaborate the group formation process.

Environment. *i*HUCOFS framework environment $E = \langle H, U, S, T \rangle$ consists of students (H), student agents (U), a teacher agent (S), and a set of tasks $T = \{T_j | j = 1, \ldots, n\}$. In this environment, the student agents works as assistants of the students and the teacher agent works with the student agents to form groups of the students to get the tasks solved by them. In *i*HUCOFS, each student agent u_i constructs and maintains a model $hm_{i,t}$ of its assigned student h_i by observing his or her behavior at time t.

Student Model. The student model in *i*HUCOFS is a two-tuple represented as: $hm_{i,t} = \langle K_{i,t}, CM_{i,t} \rangle$, where $K_{i,t}$ represents the student's *knowledge base* and

$$K_{i,t} = \{\langle ct_j, ex_{i,t} \rangle | ct_j \in T_j, ex_{i,t} \in \mathbb{R}\} \quad (1)$$

where, ct_j, describes the *area* of expertise needed to solve an assigned task T_j while $ex_{i,j,t} \in \mathbb{R}$ denotes h_i's *level* of expertise for ct_j at time t. Furthermore,

$$CM_{i,t} = \{\langle h_k, cp_{i,t} \rangle | h_k \in H \backslash h_i\} \quad (2)$$

with $cp_{i,t} \in \mathbb{R}$ represents the compatibility between users h_i and h_k as perceived by the student h_i during their collaboration to solve task T_j.

Student Groups. In *i*HUCOFS framework, we only consider non-overlapping groups and we denote the set of groups at time t as $C_t = \{C_{k,t} | \ k = 1 \ldots |C_t|\}$, where $C_{k,t} \in C_t$ at time t can be specified as:

$$C_{k,t} = \langle U_k, H_k, \ T_j \rangle \quad (3)$$

where $U_k \subseteq U, H_k \subseteq H$, and $T_j \in T$ are tasks in the environment. We also consider that at any given time *t*, each agent is a member of a group solving a particular task.

Student Group Performance Metric. In *i*HUCOFS, the performance of a student group is measured mainly by two different metrics: effectiveness and efficiency. The *effectiveness* of a group $C_{k,t}$ working on task T_j represents the quality of their solution for that task and is defined as $\xi_{k,t}$. The *efficiency* of a group is measured from the perspective of a teacher agent as the reward-to-effort ratio of its members:

$$\eta_{k,t} = r_{k,t} / oh_{k,t} \quad (4)$$

where, $r_{k,t}$ is the reward earned by group $C_{k,t}$, and and $oh_{k,t}$ is the *total* cost incurred by the members of $C_{k,t}$. For example, in a collaborative learning environment, the reward of a group could be its earned score after solving a task together and the cost could be number of messages exchanged and amount of time spent on the problem.

VALCAM Algorithm

*i*HUCOFS framework uses VALCAM (Soh et al. 2006a) – an auction-based learning enabled algorithm to form student groups. Usually, auction is used for fair distribution of some valuable resource. In *i*HUCOFS, the valuable resource is *membership in the group which has the maximum potential for improving student behavior with training*. The key idea behind VALCAM is to use the underlying auction process to form student groups that contain students who are competent and are compatible with each other. Thus, in VALCAM, the student agents bid with virtual currency that is proportional to the competence of the group being auctioned and the compatibility measure between the bidder and the members already in the auctioned group. Notice that, the competence of the members of a group is important since it allows them to: (1) engage in different types of learning scenarios (e.g., learning by teaching, learning by guiding in Inaba et al. 2000) and (2) solve the assigned problem (Teasley, S. & Roschelle 1993 as cited in Soller et al. 1999). Furthermore, balancing and assimilating students of different levels of knowledge subscribes to the *construction* view (Soller & Lesgold 2007) of collaborative learning which states that knowledge is constructed in a group by the interactions between learners of different levels of expertise. Finally, a student's willingness to collaborate with his or her group members i.e., compatibility, encourages them to work better with one another (Chalmers and Nason 2005; Issroff and Jones 2005) which would lead to improved learning and task solving.

Figures 3 and 4 describe the group formation algorithms for the teacher agent and the student agents respectively. Note that all steps of *i*HUCOFS algorithm for the teacher agent are labeled with prefix S and all steps of *i*HUCOFS algorithm for the student agent are labeled with prefix U. In *i*HUCOFS, the selected auction protocol is Vickrey (Sandholm 2000; pp. 211-219), d is the number of initial or *seed* members in the groups, and gs is the group size. The current implementation of *i*HUCOFS uses competent user first seed selection policy, i.e., puts d students with high knowledge in each student group.

S1. Initialize:
Announce Task, input: group size gs and group seed size d,

S2. Select Group Seeds:
Sort the students according to their average competence scores and choose the $d \times gs$ students

S3. Assign Group Seeds to Groups:
Assign one student seed (Step S2) for each of the groups until each group contains d users

S4. Form Group Using $VALCAM\text{-}S$:
Until all students are assigned, *do*,
For each group, *do*,
(i) Announce the group for auction to unassigned students
(ii) Accept bids from the unassigned students
(iii) Collect payment (i.e., second highest bid, cf. Vickrey) from *highest* bidder
(iv) Assign the unassigned student to the auctioned group

S5. Start Collaboration:
Announce start of collaboration to all groups

S6. Evaluate Solution and Distribute Rewards:
For each group, *do*,
(i) Input the group's task solution quality from the teacher
(ii) Provide *group* reward proportional to the solution quality
(iii) For each student, *do*,
(a) Provide individual competence score proportional to the individual contribution
(b) Provide individual reward proportional to the individual contribution

S7. Evaluate Groups:
For each group, *do*,
Calculate effectiveness and efficiency

Fig. 3. Group formation algorithm for teacher agent S.

U1. Initialize:
Initialize virtual currency balance
U2. Check Group Assignment:
If assigned to a group (*as seed*)
(i) Inform assign student
(ii) Go to Step U4
Else
(i) Post assigned human student's expertise on common Blackboard
(ii) For other students, *do,*
Post compatibility with to the common Blackboard
U3. Join Group Using *VALCAM-U*:
Until assigned to a group, *do,*
For each round of bidding for the auctioned group, *do,*
(i) Read average competence of the members from the common Blackboard
(ii) Read average compatibility of the assigned student with the auctioned group's students
(ii) Bid with an amount proportional to the average expertise U3(i) and compatibility U3(ii)
U4. Monitor h_i:
Observe student and store interaction with other students and individual effort
U5. Update Competence and Compatibility:
(i) Update assigned student's model with competence from teacher's evaluation
(ii) Update assigned student's model with compatibility from assigned student's input (survey)

Fig. 4. Group formation algorithm for student agent $u_i \in U$.

Next we describe how the VALCAM-S and VALCAM-U can be used to form groups in a typical CSCL environment. We describe the use case through four stages:

VALCAM Algorithm's Use Case:

1. **Initialization:** In a typical CSCL environment, first, the teacher would use the teacher agent to announce the task to all the students and their assigned student agents (Step S1). Furthermore, the teacher would use the teacher agent to select a number of seeds (Step S2). The seed selection allows the teacher to distribute the high-performing students in all groups to improve collaboration and knowledge building. Once the student agents receive the task announcement, they first initialize their assigned student's virtual currency account (Step U1) where the virtual currency was assigned by the teacher to the student according to his or her performance (e.g., task solution quality) in the previous CSCL sessions. If a student agent is not selected as a seed, it posts (Step U2) its estimate of the competence of the student and his or her compatibility with others to the common blackboard (which is used in later auction rounds by other student agents in Step U3(i)) and waits for the teacher agent to announce auction rounds for group formation. The competence is calculated using a weighted average of the expertise of the student (Stored in Step U5(i)) with the weights representing the similarity between a previous task and the current task. Furthermore, in Step U2, the compatibility between a student and the other students in the system are calculated using a weighted average of the assigned student h_i's previous evaluations of the other users as group members (Stored in Step U5(iii)).

2. **Auction:** Once the seeds are assigned, the teacher can signal the teacher agent to start the auction (Step S4) to allow the student agents to join the seeded groups by submitting bids. For each auction round in Step S4, the teacher agent announces an *auctioned group* to the unassigned student agents, collects their bids—where bids are posted by student agents in Step U3(ii)), selects the student agent with the highest bid to be the winning agent, and then collects a fee—based on the second highest bid (Step S4(iii))—from it. The teacher agent then officially assigns the winning agent to the auctioned group (Step S4(iv)). Once an auction round is announced by the teacher agent (Step S4(i)), the student agent uses the VALCAM-U algorithm (Step U3) to bid. According to VALCAM-U (Soh et al. 2008), the student agent reads the competence and compatibility of the users of the group being auctioned from the common blackboard (Step U3(i)) and bids a virtual currency amount proportional to the average of: (1) the *competence* of the users of the auctioned group and (2) the *compatibility* between its user and the users in the auctioned group (Step U3(ii)). While calculating the compatibility, the student agent considers its assigned student's compatibility view of the members of a group *and* those group members' compatibility views of its assigned student. After winning a bid and being assigned to a group, the student agent starts monitoring the behavior of the assigned student (Step U4) to estimate the effort toward the solution of the task. Notice that depending on the size of the group and the number of student agents, there could be a single student agent bidding in the *last* round. We assume that the student agents would always bid their true valuation regardless of the number of bidders.
3. **Collaboration:** Once the student groups are formed, the teacher can announce the start of collaboration through the teacher agent. The students can then use the underlying CSCL environment's tools and functionalities to collaborate on and solve the assigned task. While the students are collaborating, the student agents keep track of that collaboration.
4. **Evaluation:** Once the groups have completed the task, the teacher can view and evaluated the quality of the solution of the tasks completed by the student groups (Step S5(i)). For instance, in our implementation of *i*HUCOFS in I-MINDS, the teacher evaluates the solutions submitted by the groups and inputs the task solution quality (Step S5(i)) to the teacher agent's part of *i*HUCOFS algorithm. Based on the quality of the solution, the teacher agent distributes the group rewards (Step S5(ii)), the individual rewards (Step S5(iii)) in the form of virtual currency to the students. The group reward is proportional to the task solution quality (Step 5(i)) and the individual reward is calculated by multiplying the task solution quality with the ratio of individual effort (measured by the student agent in Step U4(ii)) of a student to the total effort of the members in his or her group. Once the rewards are announced by the teacher agent, the student agents update their assigned users' models: (1) by storing the individual reward provided by the teacher agent (Step U5(i)) as the expertise (i.e., competence), (2) by updating the virtual currency balance using the individual reward provided by the teacher agent in Step S5(iii) (Step U5(ii)), and (3) by storing the student's view of all other group members as compatibility (Step U5(iii)).

IMPLEMENTATION

I-MINDS

We have implemented I-MINDS using Java where the student and teacher agents are implemented as Java Objects. The teacher and student agent objects also contain an interface GUI through which their respective users are able to interact with them. During a CSCL session, I-MINDS agents communicate with each other using a KQML (Huhns & Stephens 2000)-based language. Figure 5 shows the timeline and agent communication during a typical CSCL session. The agents also have access to the common repository which is currently implemented as a set of tables in a relational database in MySQL server. Next, we provide brief overview of I-MINDS teacher agent and student agent's modules.

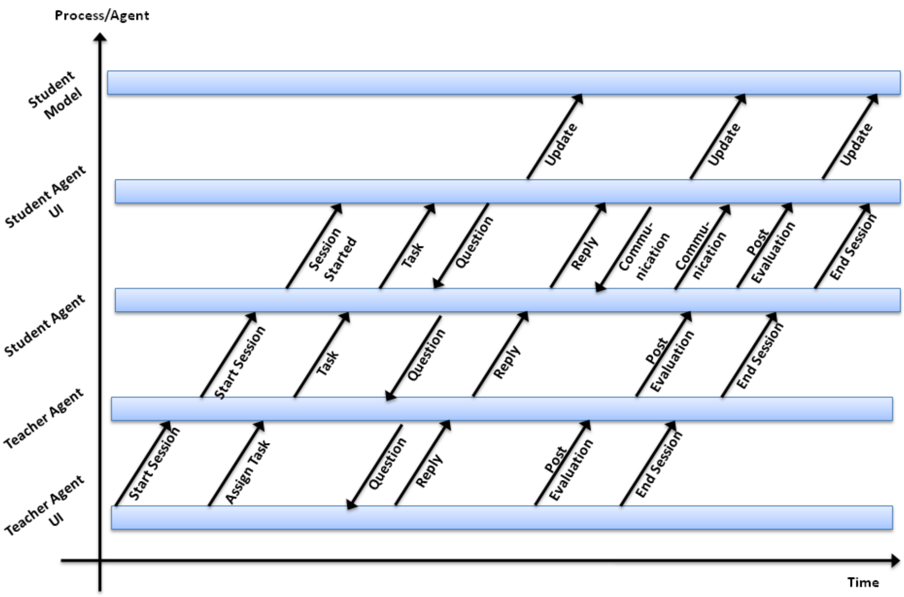

Fig. 5. Agent Communication during an I-MINDS session.

Teacher Agent Modules:
Content Dependent Module:

- **Questions and Keywords.** This module contains a reinforcement learning mechanism and uses: (1) teacher-provided ranking of questions (Soh et al. 2008) and (2) keyword stems extracted (using Porter Stemming algorithm (Jones & Willet 1997)) from task/assignment descriptions, and (3) teacher-defined rules and heuristics from the rules and heuristics module to rank the students' questions for the teacher during a CSCL session. Furthermore, this module uses the ApplePie parser (Sekine & Grishman 1995) and the utterance classifier of AutoTutor (Olney et al., 1995) to classify incoming student questions into different classes such as Contribution and Discovery.
- **Rules and Heuristics.** The rules and heuristics engine (Namala 2004) contains teacher-defined rules and heuristics that determine how the students' questions should be ranked. As the teacher responds to questions, this module adjusts the weights of the rules and heuristics (Namala 2004) to better rank the questions.
- **Classroom Profile.** The classroom profile module in teacher agent builds and maintains the profiles of all participating students in the classroom and uses them for building student groups (using *i*HUCOFS framework) and help the teacher view/grade student evaluations. Table 2 describes how and what information is collected by the teacher while maintaining the classroom profile.

Table 2: Classroom Profile of I-MINDS Teacher Agent

Category	Tracked Information	Collected by Teacher Agent from
Student Activity	Login/Logout – *Average time and frequency of login.*	Student agent
	Surveys (Appendix A) – Results of Self-Efficacy Questionnaire, Peer-Rating Questionnaire, and Team-Based Efficacy surveys	
Communication	Log of all chat and forum messages exchanged with other students during CSCL sessions	
Individual Contribution	Log of all student activities (editions, revisions, etc.) while students are collaborating.	
Performance	Group and individual student evaluations	Teacher

Content Independent Module:

- **Machine Learning.** This module contains the reinforcement learning mechanism (Namala 2004) that allows the teacher agent to update the weights of rules and heuristics and keywords.
- **Coalition Formation.** This module contains the VALCAM algorithm which allows the teacher agent to communicate with the student agents and form student coalitions for CSCL sessions.
- **Coalition Support.** Coalition support module contains the GUI for the teacher to view performance of student coalitions in terms of their individual and group grades during a CSCL session so that the teacher can intervene and help the student groups that are falling behind.
- **Repository Mechanism:** This module is implemented in Java using Jdbc connection library to allow all teacher agent modules to store and retrieve necessary information.

Student Agent Modules:

Content Dependent Module:

- **Archives:** The archive module stores a student's model and his or her performances for all previous sessions.
- **Student Profile:** This module stores the student's profile which contains information the performance of the student for the current task and session and the student's interaction with the I-MINDS agents, teacher, and other students. Table 3 shows how and what information is tracked by a student agent to build and maintain its assigned student's profile.

Table 3: Student Profile of I-MINDS Student Agent

Category	Tracked Information	Collected from	Use in ACW module and *i*HUCOFS
Student Activity	Login/Logout – *Average time and frequency of login.*	Student's interactions with I-MINDS student agent GUI	*Not Used*

	Surveys (Appendix A) – Results of Self-Efficacy Questionnaire, Peer-Rating Questionnaire, and Team-Based Efficacy surveys	Student's interactions with I-MINDS student agent's survey module	Perception attribute of ACW module (*peerRating* and *teamEfficacyRating* in Table 1)
Communication	Log of all chat and forum messages exchanged with other students during CSCL sessions	Student's interactions with Other group members through I-MINDS student agent's chat or forum tool	ACW module (Table 1) and VALCAM algorithm (Step U4(ii) in Figure 5).
Individual Contribution	Log of all student activities (editions, revisions, etc.) while students are collaborating.	Student's interactions with Topic summary revision tool of the ACW module	Student competence values and individual contribution values in VALCAM and ACW module respectively.
Performance	Group and individual student evaluations provided by the teacher for tasks/assignments.	Student agent's communication with I-MINDS teacher agent	Average individual evaluation scores for tasks is used as a measure of competence for tasks in VALCAM-S algorithm steps S2 and S6(iii)-(a) and VALCAM-U algorithm. Step U3(i) of Fig. 4

- **Repository:** The repository mechanism of I-MINDS student agent also uses a MySQL database to store and retrieve data.

Asynchronous Collaborative Writing (ACW) Module

The ACW module (Student Interface shown in Figure 6) has been implemented in the existing I-MINDS agent architecture. The assignment component of the ACW module is integrated with the I-MINDS teacher agent where the assignment module allows the I-MINDS teacher to view and or assign collaborative writing assignments. The edition component of the ACW module has been incorporated with the I-MINDS student agent where the edition component allows the students to post various types of editions (e.g., proposition and extension) to the collaborative writing assignment. The communication component of the ACW module has been implemented as chat (Figure 7(a)) and forum (Figure 7(b)) tools in the I-MINDS student agent. The approval component of the ACW module has also been implemented in the I-MINDS student agent. This approval component allows the students to view the whole collaborative writing assignment prepared by his or her group and approve it for submission to the teacher. The tracking component of the ACW module has been implemented in the background (i.e., without GUI) of the I-MINDS student agents to monitor student behavior from the communication, edition, and perception dimensions of the students.

The communication dimension of the tracking component monitors the student behavior (e.g., how many messages sent) while they are using the communication component, the edition dimension monitors student behavior (e.g., how many propositions posted) while they are using the edition component, and the perception dimension tracks: (1) *peerRating* – entered by the students through Peer-Rating Questionnaire (Soh 2004) in I-MINDS student agent and (2) *teamEfficacyRating* – entered by the students through Team-Based Efficacy Questionnaire (Soh 2004) in I-MINDS student agent.

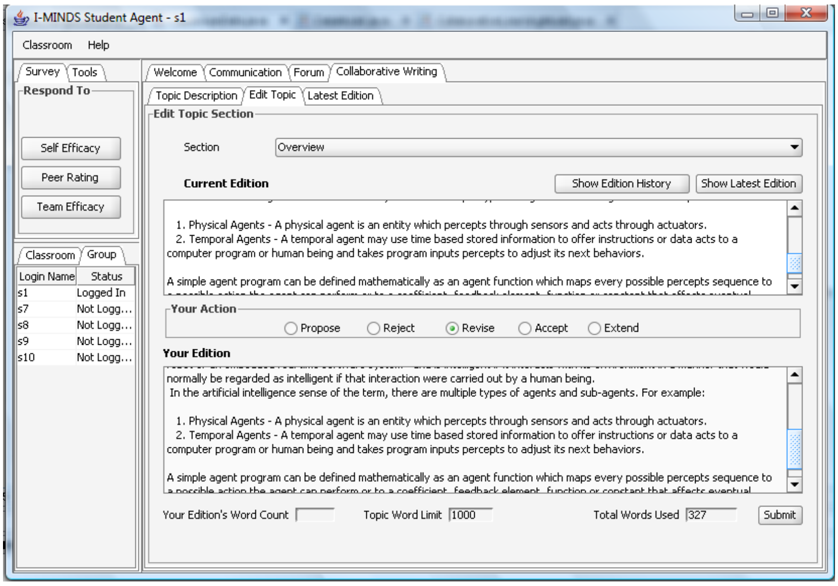

Fig. 6. I-MINDS student agent GUI showing edition component of ACW module.

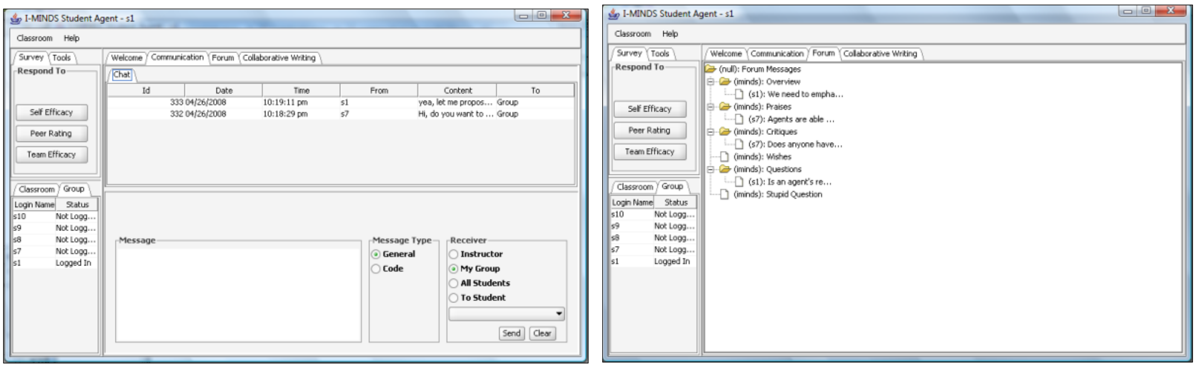

Fig. 7. (a) (left) I-MINDS student agent GUI showing chat tool, (b) (right) student forum of the communication component of ACW module.

Finally, the tracking component of the ACW module has been incorporated with the teacher agent to monitor the performance dimension (Table 1) of the students. When the teacher evaluates the collaborative writing assignments submitted by the student groups, that evaluation is stored as *groupEvaluationScore* (Table 1) by the tracking component of the ACW module. Furthermore, after every collaborative writing session, the tracking component prepares a summary (Figure 9) of the student's contribution combining the communication, edition, and perception dimensions of tracking. This summary contains: (1) the frequency count of each type of editions (e.g., propositions, accepts, revisions, extensions, rejections, and comments posted in the chat or forum) of each student, (2) the normalized (by dividing with the total for each group) frequency count of each type of editions of the students, and (3) the peer-rating received by a student. The teacher can then combine this summary with the score of the collaborative writing assignment submitted by a student's group to calculate his or her individual score which is tracked as *individualEvaluationScore* (Table 1). In our current implementation, we have used the following formula to calculate the *individualEvaluationScore* (Table 1) of the students.

Say, IG is the teacher's score assigned to the collaborative writing assignment prepared by student x's group G. Then student x's individual score is calculated based on his or her contribution counts using the following formula:

$$Student\ x's\ grade = IG \times min(1.0, N/7 \times CT(x))/P_f \quad (5)$$

where,

$$CT(x) = [w_{ps} * (PS_x/\sum_{x\in G} PS_x) + w_{re} * (RJ_x/\sum_{x\in G} RJ_x) + w_{rv} * (RV_x/\sum_{x\in G} RV_x) + w_{ex} * (EX_x/\sum_{x\in G} EX_x) + 0.5 \times w_{ac} * (AC_x/\sum_{x\in G} AC_x) + w_{cm} * (CM_x/\sum_{x\in G} CM_x) + 1.5 \times w_{pr} * (PR_x/\sum_{x\in G} PR_x)] \quad (6)$$

Here, w_{ps}, w_{re}, w_{rv}, w_{ex}, w_{cm}, w_{pr}, and w_{ac} are weights with values $w_{ps} = w_{re} = w_{rv} = w_{ex} = w_{cm} = 1$, $w_{pr} = 1.5$, and $w_{ac} = 0.5$. Furthermore, N is student x's group size, PS_x is the proposition count, RJ_x is the rejection count, RV_x is the revision count, EX_x is the extension count, AC_x is the accept count, and CM_x is the posted comment count of student x, PR_x is the peer-rating received by student x based on his or her contribution to the collaborative writing assignment, and $P_f \in [0,1]$ is the . The weight distribution of Eq. 6 allows the teacher to *specify the relative importance* of the contribution types. With the current distribution, the proposition, the revision, the rejections, the extensions and the forum comments are all weighted equally, the accept is weighted below other contributions, and the peer-rating is weighted above all other contributions. This weight distribution would motivate all students to: (1) contribute by proposing, revising, extending, rejecting, and posting comments about other's editions to improve their individual score, (2) contribute more by doing propositions, extensions, revisions, rejections, and by posting comments and less by doing accepts, and (3) not post trivial contributions (e.g., proposition and revisions) as discouraged by the subsequent lower peer-rating received from other group members.

This individual score calculation (Eq. 5 and Eq. 6) is designed to motivate the students to collaborate (e.g., proposition and extension, revision of other students' work) and to communicate (e.g., post comments in the forum) with his or her group. So, for every group, the member who collaborates and communicates the most (i.e., highest $CT(x)$) is rewarded with the highest score in the group; i.e., the score given to the final version of the writing assignment by using $P_f = 1.0$. However, for the other group members, the value of P_f is used by the teacher to determine how much they are to be penalized due to their lower levels of contributions. Depending on the nature of the collaborative writing assignment and the actual contribution counts (PS_x, RS_x, etc. in Eq. 6), the teacher can decide how much penalty a student with lower contribution should incur. If the teacher considers the writing assignment difficult and the contribution count reveals that most of the editions were done by the highest contributor, the teacher can use a high value (e.g., 0.3) for P_f which would increase the difference between their scores. On the other hand, if the teacher considers the assignment difficult and the contribution count reveals that the number of contributions posted by the highest contributor and another low contributing student are similar, the teacher may use a low value of P_f(e.g., 0.2) so that the difference between their actual scores is not too high. So, by using the weights and the factor P_f in Eq. 5 and Eq. 6, the teacher can: (1) motivate the students to do one type of contribution or another and (2) determine how the students should be rewarded or penalized for their levels of contributions for the collaborative writing assignment.

Finally, notice that our use of students' edition action counts (accept, reject, etc. in Table 1) in our formulation of Eq. 6 encourages students to contribute and collaborate more to achieve high scores. However, that could result in students gaming the system by increasing their contribution count by posting trivial editions. To prevent the students from gaming the system, we have set the weight of the peer-rating higher than the other edition counts. Since the students are able to view their group members' contribution history before they provide peer rating, a student gaming the system runs the risk of receiving low peer rating. That low peer rating would then result in low overall individual score for that student.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	I-MINDS Student Participation Report																	
2	Student Id	Group Id	# of Propo:	Normalizec	# of Reject	Normalizec	# of Revisi	Normalizec	# of Accep	Normalizec	# of Exten:	Normalizec	# of Comm	Normalizec	Normalizec	Raw Contr	Final Contributoin Score	
3	1	107	3	0.5	0	0	14	0.666667	0	0	1	0.333333	10	0.344828	0.2896	1	1	
4	2	107	2	0.333333	0	0	4	0.190476	0	0	2	0.666667	7	0.241379	0.252196	1	1	
5	3	107	1	0.166667	0	0	3	0.142857	0	0	0	0	5	0.172414	0.283933	0.518764	0.518764	
6	17	107	0	0	0	0	0	0	0	0	0	0	1	0.034483	0.17427	0.169079	0.169079	
7	5	108	0	0	0	0	1	0.142857	0	0	0	0	1	0.043478	0.245372	0.237597	0.237597	
8	6	108	2	0.333333	0	0	4	0.571429	0	0	3	1	9	0.391304	0.429303	1	1	
9	8	108	4	0.666667	0	0	2	0.285714	0	0	0	0	7	0.304348	0.325325	0.747736	0.747736	
10	11	109	0	0	0	0	13	0.382353	0	0	0	0	0	0	0.313791	0.365588	0.365588	
11	14	109	0	0	0	0	14	0.411765	0	0	0	0	0	0	0.341445	0.395971	0.395971	
12	16	109	6	1	0	0	7	0.205882	0	0	0	0	0	0	0.344764	0.738441	1	
13	9	110	2	0.4	0	0	10	0.666667	0	0	0	0	0	0	0.348651	0.681275	1	
14	10	110	3	0.6	0	0	5	0.333333	0	0	0	0	0	0	0.309628	0.599047	0.599047	
15	15	110	0	0	0	0	0	0	0	0	0	0	0	0	0.341721	0.219678	0.219678	

Fig. 8. Student contribution report generated by the tracking component of ACW module.

*i*HUCOFS Framework

We have implemented *i*HUCOFS framework for group formation in I-MINDS. The students in I-MINDS assume the role of students in *i*HUCOFS, the student agents in I-MINDS assume the role of the student agents in *i*HUCOFS and the teacher agent in I-MINDS assumes the role of the teacher agent in *i*HUCOFS. The teacher *directs* the teacher agent (using teacher agent GUI) to conduct group formation by specifying the necessary parameters such as group size gs and group seed size d, and group seed selection policy (in Step S1 of Figure 4). Each student, on the other hand, interacts with his or her student agent to form groups with other agents to collaboratively solve those writing tasks to earn scores as individual rewards. While the student groups are working, the student agents track their assigned students' effort (Step U4(ii) in Figure 5) using the tracking component (edition and communication dimension in Table 1) of the ACW module. Once the student groups complete the collaborative writing assignment, the teacher evaluates the collaborative writing assignment submitted by each group and assigns the *groupEvaluationScore* (part of the performance dimension in Table 1) to the groups. This *groupEvaluationScore* is used as the group reward in the group formation algorithm (Step S5(ii)). Then, the teacher uses the students' effort monitored by the tracking component (communication and edition dimension in Table 1) of the ACW module to assign individual student scores which are used as individual rewards in Step S5(iii) in Figure 4. The individual score of a student is further used by the student agent to update the knowledge base (Step U5(i) in Figure 5)) and to update the virtual currency balance (Step U5(ii) in Figure 5) of the student. Finally, the student agents use the tracking component of the ACW module to monitor the *peerRating* (perception dimension in Table 1) of the students to calculate the compatibility (Step U5(iii)) between its assigned student and his or her group members.

PRELIMINARY EVALUATION

For our preliminary evaluation study, we conducted a 12-week experiment in an actual classroom (a senior/graduate level course in *Multiagent Systems*). We randomly divided the 13 enrolled students in the course into two sets: *control* (7 students) and *treatment* (6 students). To verify that our random division of students did not bias our experiment, we compared the control and treatment set students' exam score in the classroom. Our analysis showed that the control set students' exam score distribution had higher average (83.88 vs 70.66) and lower standard deviation (6.62 vs. 18.21). Every week throughout the 12-week period, the teacher and the control and treatment set students carried out a collaborative writing session in the following four stages:

1) **Group Formation Stage** – the control set students were divided into two groups (3 or 4 members in a group) *randomly* while the treatment set students were divided into two groups (3 members in a group) using *i*HUCOFS' group formation algorithm (Figures and 5).
2) **Assignment Stage** – In this stage, the teacher assigned the collaborative writing assignment to the student groups using the I-MINDS teacher agent GUI. In our experiment, the collaborative writing assignments were named *topic summaries*. In the topic summaries, the students were asked to summarize a given topic taught in the classroom lecture describing its pros and cons. For our experiment, the students were given only a description of the topic summary and they had to collaborate to write the entire topic summary from scratch. An example of the assigned topic summary was the following:

   Topic Summary
   - **Topic Title**: Search algorithms for agents
   - **Sections (All sections are Required)**:
     i) an *overview* of the topic : motivations and underlying principles, etc.
     ii) a list of *praises*: a description of what you think are the important/useful aspects of the topic
     iii) a list of *critiques*: a description of what you think are the weaknesses of the topic. This critique should discuss what you think are potential limitations toward understanding and application of the concepts contained in the topic*
     iv) a list of *wishes*: what areas of the topic do you think that should be improved
     v) a list of *questions* on material that you did not understand from the lectures and textbook
   - **Hints:** What are the three general classes of search problems? How does constraint satisfaction work? What is a path-finding problem? What do you consider in a two-player game?
   - **Word Limit:** 1000
   - **Due Date:** 10/28/2010

     *Note that the teacher is not able to cover all aspects of the complex and vast topics (i.e., intelligent agents) in the limited number of lectures. Critique section is designed to further clarify the misconceptions, lack of understanding, or confusions that students may have after attending the lectures.
3) **Collaborative Writing Stage** – During this collaborative writing stage, the students could revise the assigned topic summary assigned to their group. More specifically, until the allotted time for a topic summary was over, the control and treatment set students used *identical* I-MINDS student agent GUI to (1) make changes to the assigned topic summary (Figure 8) and (2) communicate with his or her peers using the chat (Figure 7a) or the forum (Figure 7b) in the communication component. Furthermore, *any* edition of the topic summary nullified their existing approvals and *all* group members had to communicate and approve the final version again to make it ready for submission. Notice that, due to the asynchronous nature of ACW module, there were no turn-

taking protocol used and students were free to revise the sections specified by the teacher anytime. However, to prevent race conditions where group members lose their revisions by revising at the same section at the same time, we installed a *checkout* process. This checkout process allowed group members to checkout sections to edit and *warned* them if some other group member has that particular section checked out (i.e., editing) at that time. Furthermore, the students' actions (edition dimension in Table 1) were not restricted in any ways, i.e., they were free to submit revisions of any length. Once the members of a student group decided that their topic summary was ready to be submitted, they were all *required to approve* it for submission. At the end of the collaborative writing stage (once for each collaborative topic), the students evaluated their peers and their groups using Peer-Rating Questionnaire and Team-Based Efficacy Questionnaire (Soh 2004).

4) **Evaluation Stage** – In this stage, the students' submitted topic summaries were graded. To make the evaluation process unbiased, we used a *double-blind* protocol. Each student in our experiment was identified by a unique system-assigned identification number where the *identification number to student mapping* was only known to the teaching assistant. As a result, neither the course teacher who graded the topic summaries, nor the students knew whether a student belonged to the control or treatment set. Once the specified time limit for a topic summary session was over, the teaching assistant logged in to the I-MINDS teacher agent interface and printed out the final version of the collaborative writing topic summary approved by the student groups. At this point, the student groups were disbanded and the topic summary session ended. Then, the teacher evaluated the final versions of the topic summaries to determine the score for the group and calculated the individual score of the students using Eq. 5 and 6.

RESULTS

In this section, we present and discuss the empirical results collected from our experiments to investigate the impact of the use of *i*HUCOFS on the: (a) effectiveness and efficiency of the student groups, (b) perceptions of the students of their peers and their groups and (c) collaboration and learning among the students with varying competence. Furthermore, we discuss the *accuracy* and *resolution* of the ACW module in capturing the performances of the students in the topic summary. With accuracy, we estimate how *closely* the ACW module's estimate of a student's performance correlate with his or her performance in other similar tasks, leading to better student modeling. For resolution of the ACW module, we discuss how the detailed tracked information collected by the ACW module helps the teacher to identify student behavior patterns leading to: (1) insights about the usefulness of the CSCL tools that allow us to further improve the design and implementation of those tools and (2) precise intervention (i.e., intervention to reduce free-riding in our case) to improve the quality of collaboration among the students.

Impact of *i*HUCOFS on Effectiveness and Efficiency of Student Groups and Students' Perceptions

The aim of *i*HUCOFS' group formation algorithm was to form groups with compatible and competent students to improve their collaborative learning outcome. According to our definitions, effectiveness and efficiency (Eq. 4) of student groups could indicate how well the students are able to collaborate to solve the assigned problem and learn from that collaboration. To investigate the impact of *i*HUCOFS in improving the effectiveness and efficiency of the groups, we scrutinize (1) how the student's evaluation of the groups and their peers differed for the control and the treatment sets and (2) how the effectiveness and efficiency of the groups (as measured from the teacher's evaluation) *changed over time* for the control and treatment sets. Figures 9 and 10 show the average peer-rating and the average team-based efficacy values posted by the student groups in the control and treatment sets respectively.

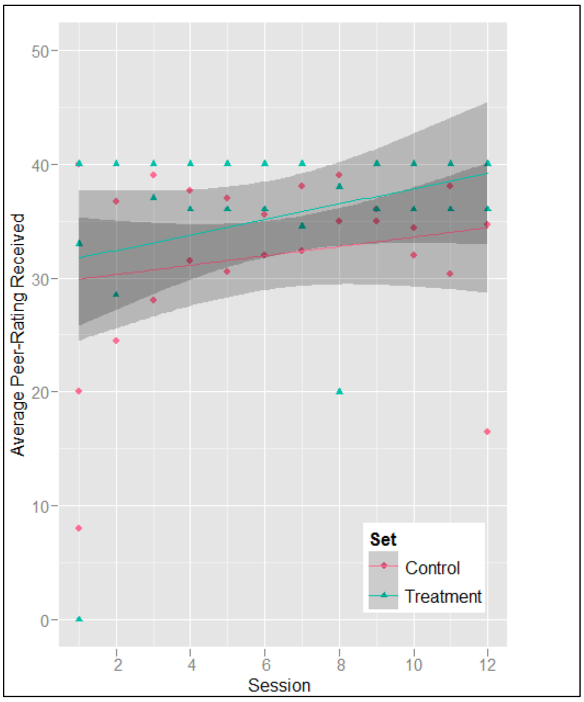

Fig. 9. Average peer-rating received by student groups in control and treatment sets.

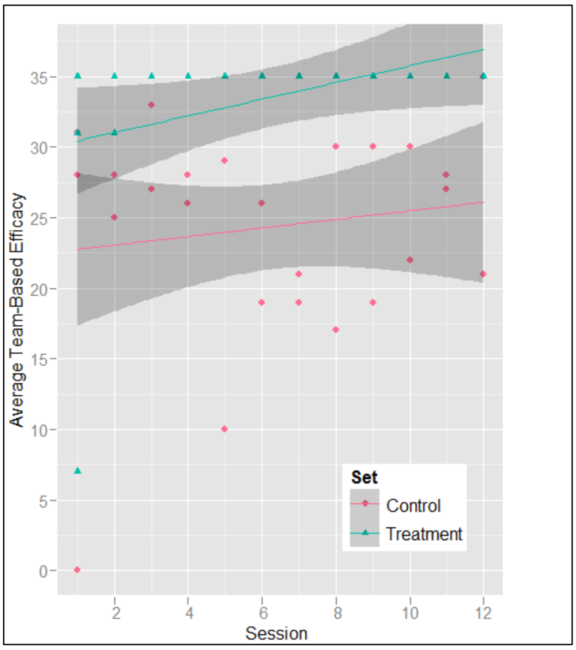

Fig. 10. Average team-based efficacy received by student groups in control and treatment sets.

Figure 9 and Figure 10 show that, although the peer-rating and team-based efficacy rating values increased for both the control and treatment sets, the treatment set students' values increased at faster rates. Though *not* statistically significant, the results indicate that how *i*HUCOFS-assigned groups could play a role, especially in improving team-based efficacy (mean rate of improvement 0.124 vs. 0.016)—the measure of how each member feels about the competence of his or her group. These relatively improved values suggest that according to the participating students, the peers and groups in the treatment set were able to learn how to work better together as a group. The students and groups in the treatment set were also able to improve their performance as a peer at a faster rate (slope of treatment set's trend line = 0.68 and slope of control set's trend line = 0.41).

Furthermore, we have calculated the effectiveness and efficiency of the groups in the control and the treatment set. The final score (IG_x in Eq. 5) of a collaborative writing assignment obtained by a group denotes how effectively they solved the assigned topic summary, i.e., effectiveness ($\xi_{k,t}$ described in the *i*HUCOFS framework section). Figure 11 shows the scatter plot and trend lines (natural cubic spline of degree ≤ 3) for the evaluation scores obtained by the groups in the control and treatment sets. The trend lines suggest that even though the treatment set coalitions had low effectiveness in the beginning, it improved around session 3 and continued at a slightly higher level. This gradual improvement, though not statistically significant, could be attributed to the treatment set students' improved knowledge in group-work (such as communication and collaboration) and task-specific expertise (such as comprehension of topics related to the course and technical writing), as a result of the balance of *compatibility* and *compatibility* of the members in the student groups. Due to *i*HUCOFS' consideration of competence and compatibility, each group contained some competent students and the students in the groups were willing to collaborate with each other. As result, the treatment set groups were able to collaborate better and were relatively more effective.

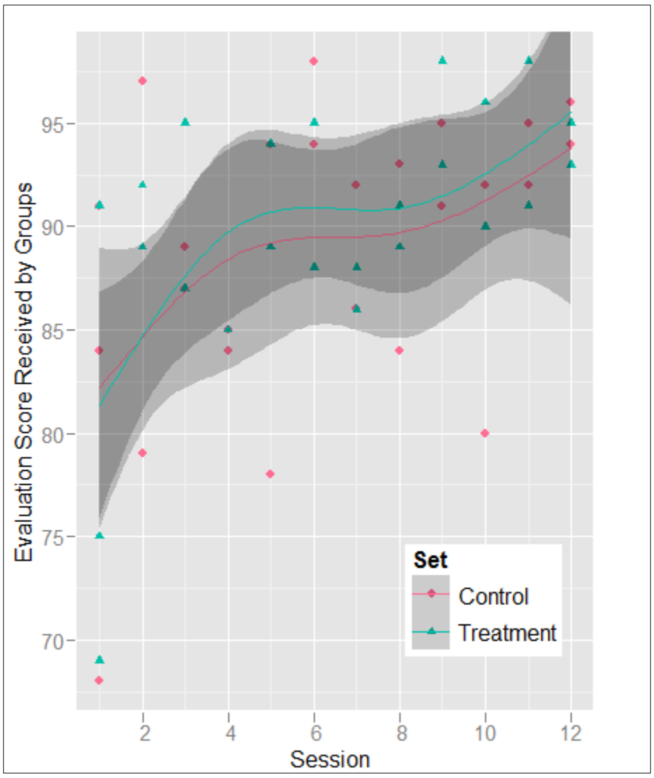

Fig. 11. Effectiveness of student groups in control and treatment sets.

Furthermore, by dividing the reward (final score IG in Eq. 5) of a coalition by the *collaborative effort* (i.e., total no. of editions and no. of messages), we calculated its efficiency ($\eta_{k,t}$ in Eq. 4). Figure 12 shows the scatter plot and trend lines for the efficiency of the coalitions in the control and treatment sets. The trend lines show that the efficiencies of the coalitions of the control and treatment set increased; however the treatment set students' efficiency values increased at a faster rate (slope of treatment set's trend line = 0.36 and slope of control set's trend line = 0.22). This could be again attributed to *i*HUCOFS' ability to facilitate learning among the treatment set coalitions by balancing competence and compatibility. Compatible coalition members are likely to be more familiar with each other's strengths and weaknesses and are able to *learn* to *coordinate* their effort better. So, the treatment set students were able to improve their efficiency better than the control set members over time.

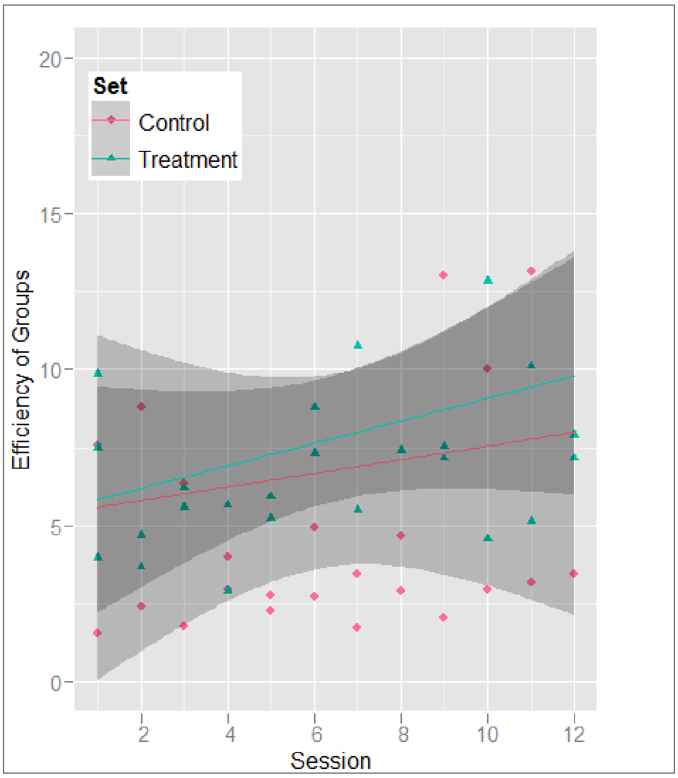

Fig. 12. Efficiency of student groups in control and treatment sets.

Impact of *i*HUCOFS on Collaboration and Learning among Students of Varying Competence

Typical CSCL groups contain members of varying competence. One possible scenario in such groups is the *sucker effect* (Roberts & McInnerny 2007): a student is perceived by his or her members to be the most able and is left to carry most of the workload. The sucker effect could reduce student collaboration since the competent students feel that they are being taken advantage of by their not-so-competent group members (Roberts & McInnerny 2007). Here we investigate: (1) whether competent students in the control and treatment sets had to *act as suckers* to complete the topic summary assignments and (2) how the students with varying levels of competence in the course contributed to the topic summary assignments (i.e., by *posting editions* or coordinating their work by *communication*).

To investigate whether any high-competence students had to act as a sucker, we look at how the overall performance (i.e., their score on the final exam) of the students in the classroom was related to the number of editions they have posted and the peer-ratings they provided to their group members. Figure 13 shows the scatter plot and trend lines for the exam scores vs. the number of editions for the students of the control and treatment sets. Furthermore, Table 4 suggests the correlations between the exam scores of a student and his or her number of editions (i.e., number of propositions and number of revisions) and peer-rating.

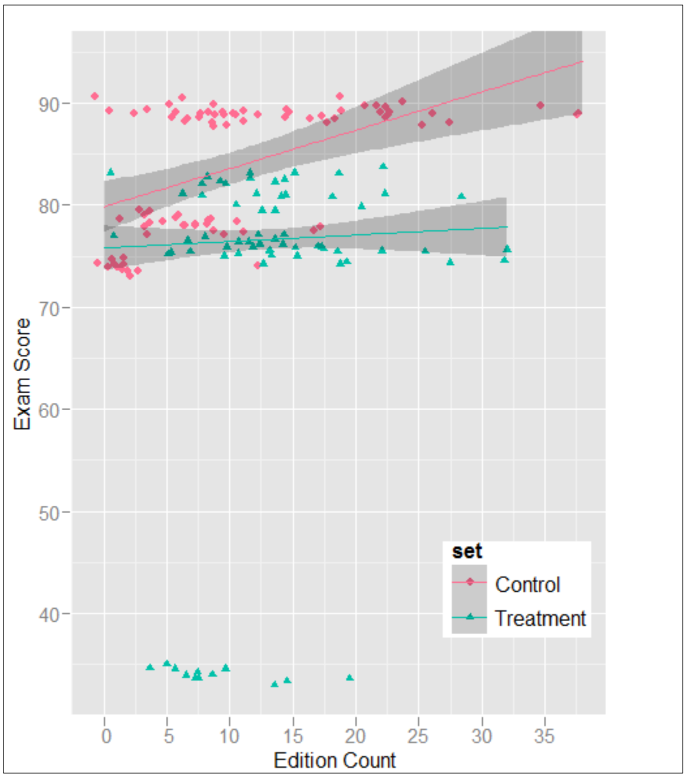

Fig. 13. Exam score vs. number of editions posted by students in control and treatment sets.

Table 4
Correlation between Tracked Collaborative Student Behavior and Other Evaluation Scores

Tracked Variable Score	Classroom Evaluation Score	Correlation	
		Control Set	Treatment Set
Total Edition Count	Exam Score	0.68	0.46
Proposition Count	Exam Score	0.61	0.41
Revision Count	Exam Score	0.59	0.59
Extension Count	Exam Score	0.78	0.42
Avg. Peer-Rating Received	Exam Score	0.38	0.50
Avg. Peer-Rating Given	Exam Score	-0.66	0.04
Average Team-Based Efficacy Rating	Exam Score	-0.77	-0.18

The trend lines in Figure 13 and the correlations in Table 4 show that the control set students who received high score in the exam (i.e., the more competent students) posted more editions, propositions, revisions, and extensions for the collaborative topic summaries. Table 4 shows that the competent control set students rated their peers and their groups *low* (-0.66 correlation between average peer-rating given and exam score and -0.77 correlation between team-based efficacy rating and exam score). In addition, the average range of the exam scores of the members of the groups in the control set was 12.57. This average range of exam scores indicates that due to random group formation, the competent and not-so-competent students were mixed together in the control set groups. So, results in Figure 13 and Table 4 hint that the *competent students* in the control set rated their *not-so-competent* *peers low*. In addition, the not-so-competent students rated their competent peers *high*. This could be explained by the not-so-competent students' inability to contribute to their respective group's effort. Since, the not-so-competent students could not contribute; their competent peers had to put in extra effort to complete the topic summary writing assignment. As a result, the competent students were not happy with the performances of their not-so-competent peers and their group as a whole. This dissatisfaction of the competent students in the control set was expressed in their low peer-rating and team-based efficacy rating.

On the other hand, for the treatment set groups, the average range of the exam score of the members of the groups was 24.13 (almost twice *higher* than the control set groups). This implies that, like the control set groups, the treatment set groups also contained both competent and not-so-competent users. However, the trend lines in Figure 13 and the correlations in Table 4 paint a different picture of collaboration for the treatment set students. The competent treatment set students who performed well in the exam also posted more editions, propositions, revisions, and extensions for the collaborative topic summaries. However, the correlations between the editions and the exam scores were *not as high as* those found for the control set students. Furthermore, Table 4 shows that the correlations between the students' performances and their evaluations of their groups and their peers were not significant (-0.04 correlation between average peer-rating given and exam score and -0.18 correlation between average team-based efficacy and exam score). These correlation values measuring the relationships between the competence of the students and their evaluations of their groups and peers indicate that, unlike the control set students, *there was no visible trend that showed the competent students' dissatisfaction with the performance of their not-so-competent peers.* The observation that the competent students performed *more* editions but were *not dissatisfied* with their not-so-competent peers could be explained by the *i*HUCOFS algorithm's effort to improve collaboration and learning of the students by forming groups by balancing competence and compatibility. This balance of competence and compatibility *promotes* various learning scenarios (e.g., learning by teaching and learning by being taught) among the students. As a result, the not-so-competent members are able to learn how to contribute more to their groups' work. Furthermore, the learning scenarios (e.g., learning by teaching, learning by being taught, and learning by discussion) require the competent students guiding and leading the not-so-competent students. The leadership activities in the topic summary context would include the competent students making frequent minor modifications to the editions posted by their not-so-competent peers. So, the competent students' relatively higher number of editions could be due to their leading or guiding activities. Furthermore, the *absence* of dissatisfaction with their not-so-competent peers indicates that either (1) the competent students were more patient with their other peers (due to their compatibility) or (2) those not-so-competent peers might be learning and collaborating and contributing to their group's effort.

Accuracy and Resolution of Student Tracking in ACW Module

Here we discuss how accurately or closely the ACW module captured the performances of the students by comparing the students' performances in the classroom (i.e., their scores in *other* classroom activities) with their performances in the topic summary (i.e., their evaluation scores). Furthermore, we discuss how the higher-resolution of information gathered by the tracking component of the ACW module helps us understand student behavior better leading to: (1) further improvement in the design and implementation of the ACW module and the scoring scheme and (2) precise teacher intervention to improve the quality of student collaboration. In this paper, we report on teacher intervention to reduce free-riding by the students.

Accuracy: Relationship between Student Performances in Topic Summary Writing and in the Classroom. The topic summary writing in the ACW module requires knowledge of the subject matter and ability to collaborate with the peers. So, one way to check the accuracy of the tracking and scoring in the ACW module is by looking at the correlation between the performance of the students in the topic summary assignments and other classroom assignments that require understanding of the subject matter and ability to collaborate. As an indicator of students' understanding of the subject matter, we can use their scores in the final exam. Furthermore, as an indicator of students' collaboration skills, we can use their scores in the *game days*. In the Multiagent Systems classroom, all students participated in several game days in which they, acting as human intelligent agents, formed groups (different from topic summary groups) and competed against each other in various scenarios such as auction and negotiation. So, good performance (i.e., high scores) in these game days required both knowledge of the subject matter and collaboration abilities. Table 5 shows the correlations.

Table 5
Correlation between Topic Summary Scores Captured with ACW Module and other Classroom Evaluation Scores

Tracked Variable Score	Classroom Evaluation	Correlation	
		Control Set	Treatment Set
Topic Summary Score	Exam Score + Game Day Score	0.58	0.55

The moderately high correlation values shown in Table 5 hint that the performance of the students captured by the tracking and scoring schemes in the ACW module closely reflects their *knowledge* of the subject matter *and* their *collaboration skills*.

Resolution 1: Understanding Student Behavior for Further Improvement of CSCL Tools and Techniques. The detailed tracking of the ACW module allows the teacher to monitor and understand student behavior better, leading to better understanding of the usefulness of the CSCL tools used in the classroom. This better understanding of the usefulness of the CSCL tools used in the classroom would allow the teacher to modify the existing tools or introduce new tools and techniques that improve the students' overall experience of the CSCL environment. Figure 14 shows how the individual student behavior regarding the topic summary editions and the evaluation scores they received changed over time across the 12 topic summary writing sessions.

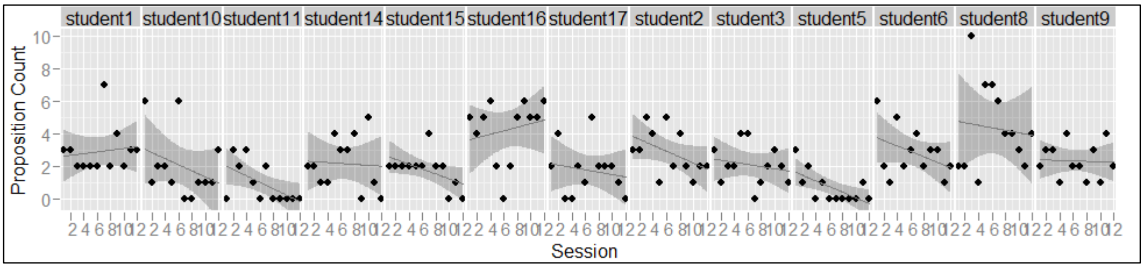

Fig. 14(a). Proposition counts of students across 12 topic summary writing sessions.

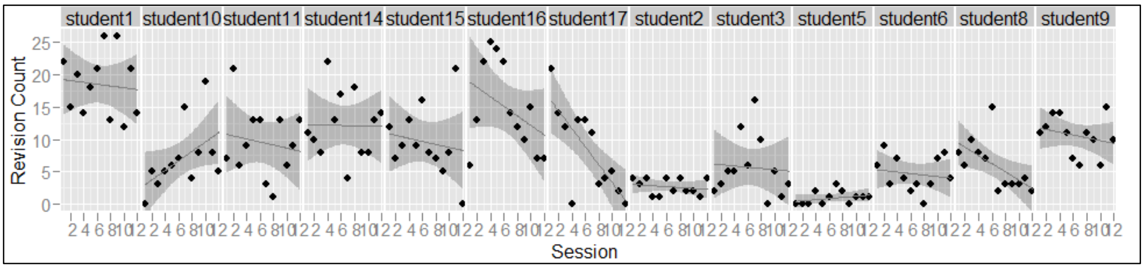

Fig. 14(b). Revision counts of students across 12 topic summary writing sessions.

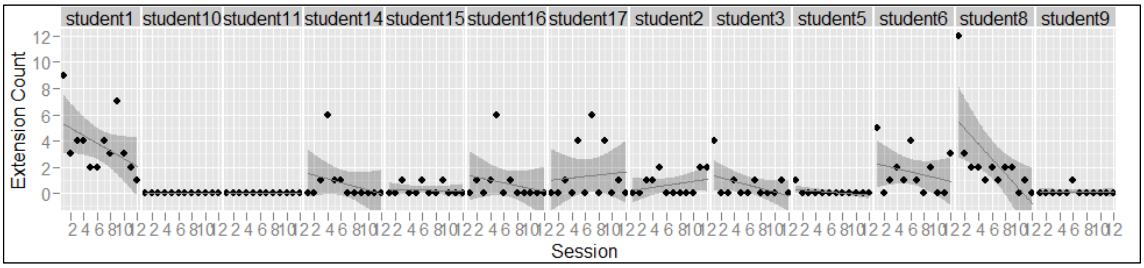

Fig. 14(c). Extension counts of students across 12 topic summary writing sessions.

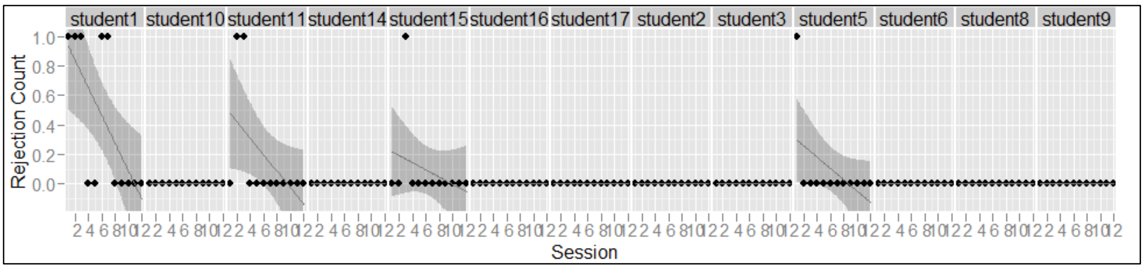

Fig. 14(d). Rejection counts of students across 12 topic summary writing sessions.

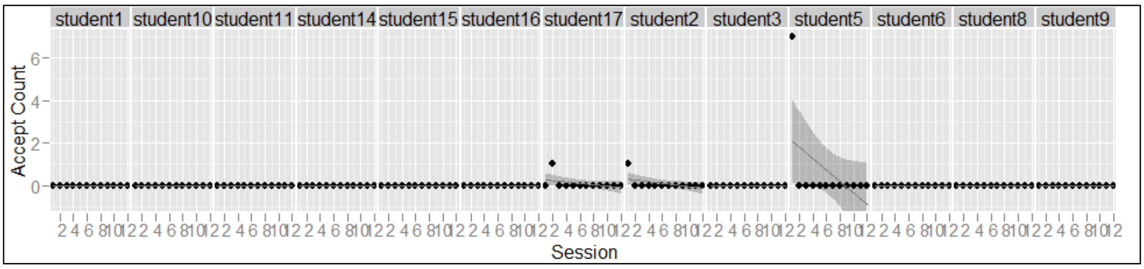

Fig. 14(e). Accept counts of students across 12 topic summary writing sessions.

Figures 14(a)-(e) show that: (1) the students mainly proposed and revised and (2) the students *rarely* extended and *almost never* rejected or accepted their while collaborating to write the topic summaries. Upon interviewing the students, we found that the students perceived the *accept* action as something that does not require any effort (no written contribution) from the students' part but allows them to increase their individual scores (according to its inclusion in Eq. 6). As a result, most of them agreed to refrain from posting *accepts*. Furthermore, the rare use of rejection of the students could be due to their inclination toward not offending their peers. In the ACW module, the most visible and direct evaluation of the students are the peer-rating and the reject actions. The peer-ratings were not visible to the students whereas the rejections were. Due to this visibility, the students could have refrained from rejecting their peer's work to avoid being seen as dismissive and to avoid potential conflicts. Finally, the extension action requires the students to accept the existing version of the topic summary and add content to it. Due to the students' lack of knowledge about the subject matter (since they have learned the topic just before they start writing the summary), the individual sections in the versions of the topic summary posted by a student were not perfect in terms of content and logical flow. As a result, the student(s) extending those versions could not first accept and then extend it. Instead, those students revised (changing content, logical flow, etc.) the already posted version and then added content to it. Based on the students' usage pattern for the *accept*, *reject*, and *extend* actions, the CSCL environment (e.g., grading and design of the ACW module) could be refined further so that the students are able to better use these functionalities. For example, the accept and the reject actions could be combined and modified to be a single anonymous polling tool that asks the students about the quality of a particular version of a section of the topic summary. The participation in this new polling tool would not count positively towards the responder (i.e., who posted the poll) but would determine the score received by the student who posted/modified the version of the polled section. This modification would: (1) allow the students to express their opinion about the existing version of the topic summary but not be perceived as someone trying to improve their scores without contributing toward their group, (2) allow the students to avoid being perceived as dismissive toward their group members, and (3) motivate the students to post better quality editions so that their peers do not reject it. Finally, the *extend* action could be disabled when the topic summary writing is implemented in classrooms where the students are newly introduced to the subject topic.

Resolution 2: Understanding Student Behavior for Facilitating Teacher Intervention—Detection of Free-Riding Students. One of the most problematic aspects of collaboration in CSCL environment is free-riders (Roberts & McInnerny 2007), i.e., students who do not contribute to the final output of the group but yet receive the same or similar rewards as those who do. We have designed an individual scoring scheme (Eq. 5) that penalizes free-riders based on the information provided by the tracking component of the ACW module to motivate the students to avoid free-riding. Our scoring scheme was successful in *discouraging* free-riding except in three different cases. Table 6 summarizes the free-riding incidents with the involved students and the teacher's action.

Table 6
Free-riding Incidents

<table>
<tr><th>Session</th><th>Student</th><th>Description</th><th>Teacher Action</th><th>Incident Repeated?</th></tr>
<tr><td>4</td><td>Student17</td><td rowspan="3">Student did not contribute and earned Low (not 0) score due to high peer-rating</td><td>Warning email sent to student</td><td>No</td></tr>
<tr><td>9</td><td>Student5</td><td>Warning email sent to student</td><td>No</td></tr>
<tr><td>12</td><td>Student15</td><td>Teacher warned student in face-to-face meeting</td><td>No</td></tr>
</table>

These three students' achievement of low scores for topic summaries with no contribution can be explained by the way our scoring scheme calculated the individual reward of the students. Our scoring scheme calculated the rewards for the students with the premise that his or her peers will provide an *objective* and *fair* assessment of their group members in the form of peer-rating. This peer-rating is combined with the other contribution counts (e.g., PS and RJ, in Eq. 5) to prepare the raw and the final individual scores of the students. Although the three mentioned students *did not contribute at all*, their group members *did not post* the lowest peer-ratings for them. As a result, those students received low scores but did not receive 0.

One way this problem could be avoided is by creating a threshold value of raw contribution scores. Any student who does not achieve that minimum threshold of raw contribution score (i.e., a student who does not contribute enough) would receive a 0 for that topic summary. However, there could be situations where a student cannot reach that threshold since his or her group members have already written a good enough topic summary leaving less room for improvement. Some precautions could be taken to mitigate such scenarios. For example, the minimum contribution threshold could be set adaptively by taking into the performance of the students who have already contributed. If the students who have contributed are students who have been deemed to be knowledgeable and skilled in writing high quality topic summaries in the past, then it is natural to expect the amount of changes by other group members to the writeup of that competent student would be low, and vice versa.

Summary

To summarize, the analysis of the data in our experiment suggest that, the use of *i*HUCOFS for group formation may: (a) improve the effectiveness and efficiency of the student groups, (b) improve the perceptions of the students of their peers and their groups, and (c) improve collaboration among students with low and high competence. Furthermore, our discussions regarding the *accuracy* of the ACW module indicate that the performance of the students calculated from the tracked information correlate well with their performances in other similar classroom activities. Finally, our discussions regarding the resolution of the ACW module hint that the detailed tracked information collected by the ACW module helps the teacher to identify student behavior patterns leading to: (1) insights into the usefulness of the CSCL tools (e.g., extend, reject and accept actions in the ACW module) that allow us to further improve the design and implementation of those tools and (2) *precise* intervention (i.e., intervention to reduce free-riding in our case) to improve the quality of collaboration among the students.

RELATED WORK

To compare our research effort with the current state of the art of group formation and tracking and evaluation methods in CSCL systems, we have divided our related work section into two sub-sections. First, we discuss the differences between the *i*HUCOFS framework with other CSCL group formation methodologies. Then we discuss recently developed CSCL systems and systems that support collaborative work to describe how they track and evaluate student performances. For each subsection, we first describe the methodologies before summarizing their relations to the *i*HUCOFS framework.

Group Formation in CSCL Systems

To avoid complications arising from allowing the students choose their own groups (e.g., loss of heterogeneity and lack of expertise in a group), Redmond (2001) proposed a group formation algorithm that forms learner groups for participating in projects by gathering students who do not have conflicting schedules. Although their group formation program could generate student groups whose members were able to collaborate without any schedule conflicts, sometimes the membership of the groups had to be adjusted by the teacher.

To promote heterogeneity in the student groups, Graf and Bekele (2006) used Ant-Colony Maximization process to form student groups. Their system formed student groups that contained low, average, and high scoring students. They modeled the group formation problem as a directed graph and used ACS (Ant Colony System) algorithm to form groups to optimize the heterogeneity of the groups. The authors reported that their group formation algorithm achieved near-optimal solutions (in terms of heterogeneity of the groups) for 100 and 512 students.

Muhlenbrock (2006) discussed forming learner groups based on information from learner profile (e.g., what he/she knows, what his/her area of difficulty is, where he/she is at) and learner context (e.g., when he/she is available). The authors discussed that by forming groups based on these two factors, they could improve the quality of formed groups since it allows for the ad-hoc creation of learning groups useful for providing peer help for immediate problems. The authors tested their group formation framework with a set of experiments and developed a distributed application that helps teachers to form learning groups.

Christodoulopoulos et al. (2007) proposed Omadogenesis, a web-based group formation tool to support the teacher to automatically create homogeneous and heterogeneous groups based on up to *three* (e.g., knowledge, gender, learning style) criteria. Furthermore, in their proposed group formation method, the students were allowed to negotiate their grouping and the teacher was able to manually adjust the formed groups. The authors proposed method formed heterogeneous groups by combining students with *low*, *average,* and *high scores* using the heterogeneity matrix and formed homogeneous groups using the Fuzzy C-Means. Their preliminary results indicated that their tool could form heterogeneous and homogeneous groups. However, their results did not describe any studies that compared the improvement of learning of the students in the groups formed using their method against those of students in any other (e.g., random) group formation method.

Wang et al. (2007) used the Random Mutation Hill Climbing (RHMC) to design DIANA – a group formation algorithm to form heterogeneous student groups. The authors designed the DIANA grouping system to form groups to achieve: fairness (in the form of groups having the same size), equity (assigning all students to their most suitable group), flexibility (allowing teachers to address single or multiple psychological variable), and heterogeneity (guaranteeing individual diversity for promoting intra-group interactions). In their experiment, the authors compared the performances of student groups formed by DIANA (based on their thinking styles) and groups formed randomly. The results of the authors' experiment showed that both types (random and DIANA-formed) of groups were equally capable of correctly completing the assigned tasks, but the DIANA-assigned groups correctly completed a significantly larger percentage of tasks and showed less inter-group performance variance.

Summary: Although these group formation methods differ in their approaches, there are some common differences between these approaches and our *i*HUCOFS framework for group formation in I-MINDS. For example, the mentioned group formation methods do not track, model, and utilize the students' *preferences of group members*. However, recent CSCL research (Chalmers & Nason 2005) suggests that social relationship and students' preference of group members play a role in determining how well they work as a group. Furthermore, some of the mentioned group formation methods do not track, model, and utilize the *different knowledge levels of the students* which could help their group formation method to adapt to the students' learning and changing behaviors to form better groups over time. Finally, some of the group formation methods (e.g., (Christodoulopoulos et al. 2007) and (Wang et al. 2007) use psychometric classification (e.g., thinking styles and learning styles) of the students to form heterogeneous or homogeneous groups. However, the *validity* and *reliability* of psychometric tests often vary among student sets that differ in their attributes such as age and background knowledge. As a result, not all psychometric tests can accurately classify the students in a typical CSCL classroom. Therefore, the use of psychometric tests in CSCL group formation may limit the use of the group formation method to students with a particular set of attributes. *i*HUCOFS framework, unlike these mentioned research approaches, uses the knowledge and compatibility of the students to form student groups to improve students' learning. Our preliminary results suggest that *i*HUCOFS is able to form effective and efficient student groups that also improve students' perceptions about their groups.

Tracking and Assessment in Recent CSCL Systems

I-Help (Vassileva et al. 2002; Bull et al. 2001; Vassileva et al. 1999) is built on a multiagent architecture that combined a one-to-one network and a discussion forum to provide offline peer help to learners. Every learner in I-Help was represented by an agent who modeled his or her knowledge and behavior. When a learner sought help, his or her representative agent communicated with the other agents in the system and found the most suitable learner who could provide peer help. Although I-HELP tracked the users' performances in providing help using virtual currency payments, it did not track or assess the effectiveness of the negotiation process or help sessions among users.

Teixeira et al. (2002) presented MATHNET, a multiagent CSCL environment where the students could learn by interacting and collaborating with the system and among themselves. MATHNET facilitated collaborative learning with tutor agents, pedagogical agents, and learner modeling agents. The learner modeling agents modeled the learners, the searching agents selected the appropriate learning material the learners, the pedagogical agents and the tutoring agents provided the appropriate teaching strategy for the CSCL session. For the teacher, the MATHNET provided the capability of monitoring and evaluating individual as well as group activities. For learners, MATHNET provided tools that the learners could use to communicate with the system, their peers, their own group, and other groups by exchanging messages.

Constantio-González et al. (2003) proposed a web-based environment called Collaborative Learning Environment for Entity-Relationship Modeling (COLER) in which students could solve Entity-Relationship (ER) problems while working synchronously in small groups at a distance. COLER, like I-MINDS, included a message exchange tool and a common digital whiteboard where the students worked collaboratively. Furthermore, COLER allowed the teacher to form groups, to monitor and evaluate the individual and collaborative work of the learners during and after collaboration. Finally, COLER used coaches to monitor and evaluate the learners' collaborative performances as well as the performance of the groups.

Soller and Lesgold (2007) (and also Soller et al. (2003)) discussed how Hidden Markov Models (HMM) could be used to analyze online knowledge sharing interactions. The knowledge sharing episode is defined as a segment of interaction in which one student attempts to present or explain new knowledge to his peers while the peers try to understand and assimilate that new information. The researchers collected and categorized the student knowledge sharing interactions while they collaboratively solved object-oriented design problems using a chat interface. The researchers classified the knowledge sharing episodes as *effective* or *breakdown* episodes and for each episode, they identified the main topic of conversation. The researchers then used this data to train two 6-state HMMs and then used those HMMs to identify the most likely class of a new knowledge sharing episode. This analysis of knowledge sharing is relevant to our research since it may help the teachers identify the knowledge sharers in the classroom and (1) encourage effective knowledge sharing by rewarding the students according to the effective knowledge sharing episodes and (2) distribute those knowledge sharers into different groups to improve those groups' knowledge co-construction process.

Stevens et al. (2004) described how they have developed predictive models of students' learning of problem-solving skills in a general qualitative chemistry course. The researchers used self-organizing artificial neural networks to identify the most common student strategies of learning while they were working on the online tasks. The researchers then applied HMMs to the sequences of those strategies to model the learning trajectories. The researchers have found that these models and trajectories can be used to predict future performances and strategies of the students with high accuracy (>80%). Their research provides interesting possibilities for the researchers who could, based on the derived model and predictions, (1) determine whether or not the student is likely to need help in the near future and (2) strategically construct collaborative learning groups containing heterogeneous combinations of various behaviors such that the need for intervention by a human teacher can be lessened.

Gogoulou et al. (2005) presented ACT—a web-based adaptive communication tool. ACT (developed as a component of the SCALE system (Grigoriadou et al. 2004)) supported and guided the learners' communication/collaboration by implementing the structured dialogue through sentence openers or communication acts. The ACT tool aimed to guide and support the learners appropriately to: (i) eliminate the off-task discussions, (ii) guide the learners towards the underlying learning outcomes of the activity or the duties and responsibilities implied by the model of collaboration, and (iii) enable the automatic interpretation of the learners' interaction as well as the tracing of the dialogue states. Their results showed that the learners found the ACT tool useful for collaboration.

Erkens et al. (2005) described TC3—a groupware environment that allows the students to write argumentative essays collaboratively. For collaborative writing, TC3 provided the students: (1) access to relevant information regarding their essays, (2) a private notepad, (3) a chat facility including a chat history, (4) a shared work processor, and (5) planning tools for writing (a shared argumentation diagram drawing tool and a shared outline tool). In their experiment, the researchers allowed a set of high-school students to use TC3 to write several essays and then analyzed: (1) their discussions, (2) their pattern of activities during discussion and collaboration, (3) their contributions (searching for information, preparing the outline of the essay, discussing with peers, etc.) during the various phases of writing. The results of their qualitative experiment showed that the tools provided in the TC3 system helped the students collaborate and coordinate their actions regarding the planning and writing of the assigned essays. However, the authors did not use of the tracked information to evaluate the performance of the students, to improve the group formation process, and to improve the collaboration of the students (e.g., intervention to reduce free-riding).

Israel (2007) described an Intelligent Collaborative Support System (ICSS) that supported collaborative effort of students by analyzing the collaborative process dynamically while employing a web-based interface. The primary goal of ICSS was to assist members of a group to more effectively collaborate in solving a problem especially when they are working at a distance. The ICSS also provided support for students to learn the collaborative skills needed for a distributed work environment. For example, ICSS provided support for the discussion skills by examination of sentence openers chosen from a menu, keywords found in free-text sentence closers, student and group models, and historical database files. ICSS also assessed the performances of the groups by categorizing the statements of the group members and monitoring creative conflicts. Results of qualitative experiments showed that students found the system useful and that the system could guide students to work more effectively thereby making the groups more productive.

Summary. Although the aforementioned CSCL environments were designed to provide support for collaboration to the students, they do not address the issue of tracking and assessment of the performance of the students as an individual or as a group member by the teachers. However, as CSCL researchers (e.g., Roberts & McInnerney (2007)) suggested, assessment or evaluation of the students as group members is one of most important issues in the CSCL research. In I-MINDS, an instantiation of *i*HUCOFS framework, we have tried to address the assessment of students using the tracking component of the ACW module which monitors and assesses the contributions of students to their group using the students' actions, their peer's evaluations, and the teacher's evaluation of the final output of their group. The results of our experiment show that our assessment method: (1) captures student performances that closely represent the students' performances in other similar tasks and (2) provides sufficient resolution that allows the teacher to detect student behavior patterns to: (a) further improve the design of the CSCL environment and (b) provide precise intervention to improve the quality of collaboration among students.

CONCLUSIONS

In this paper, we have discussed how *i*HUCOFS' group formation and the newly designed ACW module of I-MINDS was used to address two shortcomings of typical CSCL systems, i.e., formation of student groups that improve the learning outcome of the students and assessment of a student's contribution to his or her group. The results of our semester-long experiment suggest that the *i*HUCOFS framework can utilize the tracked information provided by the ACW module to form student groups that: (a) become more effective and efficient over time, (b) improve the perceptions of the students of their peers and their groups, and (c) improve collaboration among students with low and high competence. Furthermore, our discussions regarding the tracking and assessment capability of the ACW module indicate that the detailed tracking capabilities of the ACW module provides information to the teacher that allows him or her to better understand student behavior leading to: (1) improvement of the design of the CSCL tools (e.g., the ACW module itself) and (2) precise intervention to improve collaboration among the students (i.e., intervention to discourage free-riding among students in our case).

In our future work, we aim to overcome the following limitations of our current implementations and our preliminary evaluation in the following three categories.

Qualitative Student Evaluation. We would like to incorporate more qualitative aspects of student editions while assessing a student's contributions using ACW module. We would like to allow students to rate their group member's editions during the collaborative session and use that rating to calculate the individual score. Furthermore, we are now implementing information retrieval mechanisms to determine how much of a student's contribution (i.e., words or sentences added) survives his or her group members' editions (i.e., of good quality) and make it to the final submitted version.

Improving Task Performance. We would like to incorporate tools in *i*HUCOFS which would help students improve their task performance. For example, we plan to encourage student knowledge sharing by identifying and rewarding students who are helping their group members learn (i.e., acting as the knowledge sharer) using hidden Markov models (similar to Soller & Lesgold 2007). Furthermore, we are incorporating information retrieval tools to provide topic-specific information while they are collaborating to improve their competence on the topic and thus help them improve their task performance.

Significant Result. We are also planning a comprehensive and large-scale experiment to obtain more significant results on the impact of *i*HUCOFS framework and ACW module on the group formation and individual evaluation of an asynchronous CSCL classroom.

ACKNOWLEDGEMENTS

The authors would like to thank the NSF (SBIR grant DMI-044129) and Microsoft Research ConferenceXP (CX) for the research funding. The authors would also like to thank Adam Eck and Lee Dee Miller for their help in deploying I-MINDS.

REFERENCES

Bull, S., Greer, J., McCella, G., Kettel, L. & Bows, J. (2001). User modeling in I-Help: what, why, when and how. *Proc. 8th Int. Conf. on User Modeling* (pp. 117–126).

Chalmers, C. & Nason, R. (2005). Group metacognition in a computer-supported collaborative learning environment. *Proc. Int. Conf. on Computer Education (ICCE'05)* (pp. 35-41).

Christodoulopoulos, C.E. & Papanikolaou, K.A. (2007). A group formation tool in an e-learning context. *Tools with Artificial Intelligence*, 2, 117-123.

Constantio-González, M. A., Suthers, D. D., & Escamilla de los Santos, J. G. (2003). Coaching web-based collaborative learning based on problem solution differences and participation. *International Journal of Artificial Intelligence in Education*, 13, 263-299.

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, 21, 463-486.

Gogoulou, A., Gouli, E., Grigoriadou, M., & Samarakou, M. (2005). ACT: a web-based adaptive communication tool. *Proc. Computer Support for Collaborative Learning 2005: the next 10 years!* (pp.180-189).

Graf, S. & Bekele, R. (2006). Forming Heterogeneous Groups for Intelligent Collaborative Learning Systems with Ant Colony Optimization. *Intelligent Tutoring Systems*, 4053/2006, 217-226.

Grigoriadou, M., Gogoulou, A., Gouli, E., & Samarakou, M. (2004) The activity as a fundamental unit of learning + collaboration in SCALE. *Proc. 4th Hellenic Conference with International Participation* (pp. 525-534).

Huhns, N. M., & Stephens, L. M. (2000). Multiagent systems and societies of agents. In G. Weiss (Ed.) Multiagent Systems: a Modern Approach to Distributed Artificial Intelligence (Chapter 2, pp. 88-128). Cambridge, MA: MIT.

Inaba, A., Supnithi, T., Ikeda, M., Mizoguchi, R., & Toyoda, J. (2000a). How can we form effective collaborative learning groups? *Proc. Int. Conf. on Intelligent Tutoring Systems (ITS)* (pp. 282-291).

Israel, J. (2007). Supporting collaborative learning with an intelligent web-based system. *International Journal of Artificial Intelligence in Education*, 17, 3-40.

Issroff, K. & Jones, A. (2005). Learning technologies: affective and social issues in computer-supported collaborative learning, *Computers and Education*, 44, 395-408.

Johnson, D. W. & Johnson, R. (1994). *Leading the cooperative school* (2nd Ed.). Edina, MN: Interaction Book Company.

Jones, K. S. & Willet, P. (1997). *Readings in Information Retrieval* (p. 454). San Francisco: Morgan Kaufmann.

Khandaker, N. & Soh, L-K. (2008). Forming and Scaffolding Human Coalitions: A Framework and an Implementation for Computer-Supported Collaborative Learning Environment. *International Transactions on Systems Science and Applications*, 4(1), 3-22.

Liu, X., Zhang, X., Soh, L-K., Al-Jaroodu, J., and Jiang, H. (2003). A distributed multiagent infrastructure for real-time, virtual classroom. *Proc. Int. Conf. on Computer Education (ICCE'03)* (pp. 640-647).

Muehlenbrock, M. (2006). Learner group formation based on learner profile and context. *International Journal of E-Learning,* 5(1), 19-24.

Namala, S. (2004). *An intelligent module for I-MINDS*. M.S. Project Report, Computer Science & Engineering, University of Nebraska, Lincoln, NE.

Olney, A., Louwerse, M., Mathews, E., Marineau, J., Hite-Mitchell, H., & Graesser, A. (1995). Utterance classification in Autotutor. *Proc. HLT-NAACL Workshop* (pp. 1-8). Philadelphia: Association for Computational Linguistics.

Redmond, M. A. (2001). A Computer Program to Aid Assignment of Student Project Groups. *ACM SIGCSE Bulletin*, 33(1), 134-138.

Roberts, T. S. & McInnerney, J. M. (2007). Seven problems of online group learning (and their solutions). *Educational Technology & Society*, 10(4), 257-268.Sandholm. T. (2000). Auctions, in Weiss, G. (Eds.) *Multiagent Systems: A Modern approach to Distributed Artificial Intelligence,* Chapter 5, MIT Press, 211-219.

Sekine, S., & Grishman, R. (1995). A corpus-based probabilistic grammar with only two non-terminals. Proc. 4th International Workshop on Parsing Technologies (pp. 216-223).

Soh, L-K., & Khandaker, N. (2007). Formation and scaffolding human coalitions with a multi-agent framework. *Proc. Autonomous Agents and Multiagent Systems (AAMAS'07)* (pp. 394-396).

Soh, L-K. (2004). On Cooperative Learning Teams for Multiagent Team Formation, in Technical Report WS-04-06 of the AAAI's 2004 Workshop on Forming and Maintaining Coalitions and Teams in Adaptive Multiagent Systems, San Jose, CA, pp. 37-44.

Soh, L.-K., Khandaker, N., & Jiang, H. (2008). I-MINDS: A Multiagent System for Intelligent Computer-Supported Collaborative Learning and Classroom Management. *International Journal of Artificial Intelligence in Education*, 18(2), 119-151.

Soh, L.-K., Khandaker, N., & Jiang, H. (2006a). Multiagent coalition formation for computer-supported cooperative learning. *Proc. Innovative Applications of Artificial Intelligence (IAAI'06)* (pp. 1844-1851).

Soh, L.-K., Khandaker, N., Liu, X., & Jiang, H. (2006b). A computer-supported cooperative learning system with multiagent intelligence. *Proc. Int. Conference on Autonomous Agents and Artificial Intelligent Systems (AAMAS)* (pp. 1556-1536).

Soh, L.-K., Liu, X., Zhang, X., Al-Jaroodi, J., Jiang, H., & Vermuri, P. (2004). I-MINDS: an agent-oriented information system for applications in education. *Lecture Notes in Artificial Intelligence (LNAI 3030) Series Agent-Oriented Information Systems*, Springer-Verlag, 16-31.

Soller, A. & Lesgold, A. (2007). Modeling the process of collaborative learning. Hoppe, Ulrigh H., Ogata, H., Soller, A. eds. *The Role of Technology in CSCL* (pp. 62-86). New York, NY:Springer.

Soller, A. & Lesgold, A. (2003). A computational approach to analyzing online knowledge sharing interaction. Hoppe, U., Verdejo, F. and Kay, J. (eds.) *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* (pp. 253-260). Amsterdam, the Netherlands: IOS Press.

Soller, A., Lesgold, A., Linton, F., & Goodman, B. (1999). What makes peer interaction effective? Modeling effective communication in an effective CSCL. *Proc. 1999 AAAI Fall Symposium: Psychological Models of Communication in Collaborative Systems* (pp. 116-123).

Stevens, R., Soller, A., Cooper, M., & Sprang, M. (2004). Modeling the development of problem solving skills in chemistry with a web-based tutor. *Lecture Notes in Computer Science*, 58-591.

Teasley, S. & Roschelle, J. (1993). Constructing a joint problem space. In S. Lajoie and S. Derry eds. *Computers as cognitive tools* (pp. 229-257). Hillsdale, NJ: Lawrence Erlbaum.

Teixeira, C. Labid, S., & Nascimento, E. (2002). Modeling the cooperative learner based on it's actions and interactions within a teaching-learning session. *Frontiers in Education,* 1, 19-23.

Vassileva, J., McCalla, G., & Greer, J. (2002). Multi-agent multi-user modeling in I-Help. *User Modeling and User-Adapted Interaction: Special Issue on User Modeling and Intelligent Agents*, 13(1), 1-31.

Vassileva, J., Greer, J., McCalla, G., Deters, R., Zapata, D., Mudgal, C., & Grant, S. A. (1999). A Multiagent design of a peer-help environment, in Lajoie, S. & Vivet, M. (eds), *Artificial Intelligence in Education*, 38-45, IOS Press.

Wang, D.Y., Lin, S.S.J., & Sun, C.T. (2007). DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior*, 23(4), 1997-2010.

Wooldridge M. (2000). Intelligent Agents. In G. Weiss (Ed.) Multiagent Systems: a Modern Approach to Distributed Artificial Intelligence (Chapter 1, pp. 48-54). Cambridge, MA: MIT.

Appendix A

Questionnaires Administered by I-MINDS Student Agent

A.1 Self-Efficacy Questionnaire

Please answer the following survey using one of the following answer choices:

Answer Choices: Strongly Disagree, Disagree, Neither Agree or Disagree, and Agree, Strongly Agree
1. I am experienced in working as a team member
2. I like to participate in teams
3. I have had positive experience in working in teams in this environment
4. I would rather work in a team than on my own
5. I am highly motivated to make this team successful
6. I expect the team to be very successful in accomplishing the required outcomes
7. I expect my personal contribution to be significant in the team's outcomes
8. I feel that the contributions of the team members will be equally significant
9. I am concerned about majority of the points being tied to team

A.2 Peer-Rating Questionnaire

Please answer the following survey using one of the following answer choices:

Answer Choices: Strongly Disagree, Disagree, Neither Agree or Disagree, and Agree, Strongly Agree

1. The group member has a sharing attitude toward other team members
2. The group member has a positive attitude toward the team
3. The group member has been truly earning the rewards he/she has received
4. The group member is willing to help other team members anytime
5. The group member eagerly accepts and shares all team responsibilities
6. The group member attempts to accomplish team's missions and goals
7. The group member participated in establishing the teams mission and goals
8. The group member participated in team discussions
9. The group member's level of contribution to the team (0-100)

A.3 Team-Based Efficacy Questionnaire

Please answer the following survey using one of the following answer choices:

Answer Choices: Strongly Disagree, Disagree, Neither Agree or Disagree, and Agree, Strongly Agree

1. Over the course of the team work, our team was successful in working together as a team
2. Over the course of the teamwork our team was successful in solving conflicts within our team
3. Over the course of the teamwork, our team had little problem with

4. As the teamwork process draws to a close, I feel more comfortable having
5. I believe that working in the team will be a valuable experience for me
6. I would like to participate as a team member in the future
7. Cooperative teams should continue to be a required element of this environment
8. Denote the percentage of the work done by your team was done by each of your team members.