International Journal of Continuing Engineering Education and Lifelong Learning, Vol. 12, No. 1-4, pp. 149-158, 2002.

Learning by Judging: A Network Learning Environment Based on Peer Evaluation*

Chuen-Tsai Sun

Department of Computer and Information Science
National Chiao Tung University, Taiwan
ctsun@cis.nctu.edu.tw

Abstract

The paradigms of science and engineering education have been evolving along the dimensions of theory, pedagogy, and technology. As network-based computer-assisted learning is drawing more and more attention in the past few years, promising learning theories such as constructive learning and collaborative learning have found their new meaning and new ways of implementation in the new, network-based environment. In this paper, we propose a learning environment in which a learning strategy, called Learning by Judging, can be realized to help constructive learning.

The proposed environment is a World Wide Web (WWW) based system that supports learning strategies of learning through design and peer evaluation. Samples designed by students are demonstrated in screen windows for others to evaluate via network. A learning procedure includes several rounds of sample-design and mutual evaluation. The learning process is recorded and analyzed and then the result is fed back to the students to achieve the goal of meta-cognition. We expect this approach to encourage students participating in instructional activities early and effectively.

We conducted a small-scaled experiment in which color-matching was chosen as the design subject. We collected and analyzed users' behavior patterns in a preliminary manner. We are improving the functionality and interface of the learning environment. We expect to conduct formal instructional experiments in the near future.

Keywords

Distance Learning, Cooperative Learning, Learning Strategy, Peer Evaluation, Learning through Design.

Biographical Note

Chuen-Tsai Sun received his B.S. degree in electrical engineering and his M.A. degree in history, both from National Taiwan University, Taiwan. He received his Ph.D. degree in computer science from the University of California at Berkeley in 1992. From 1991 to 1992 he was with the Lawrence Livermore National Laboratory, California, where he participated in research projects on fuzzy neural networks. Since 1992 he has been with National Chiao

* This work was sponsored by the National Science Council, Taiwan, Project Number: NSC 86-2511-S-009-004ICL.

Tung University, Taiwan, where currently he is a Professor in the Department of Computer and Information Science. He is engaged in research and teaching in the areas of neuro-fuzzy systems, evolutionary computing, distance education, and computer-assisted learning.

I. Introduction

Students were largely isolated in traditional computer-assisted learning (CAL) environments, e.g., a CD-ROM courseware-based system, because usually the interactions provided by this kind of environments occurred only between the students and the courseware/system. Recently, the development of computer networks has enhanced the CAL environment with communication facilities among students in a synchronous or asynchronous manner so that students can cooperate with each other in a virtual learning community. There have been many network-based platforms developed and evaluated, among them the World Wide Web has become one of teachers' and educational researchers' favorites.

This trend enables the exploration of new pedagogical strategies in computer-assisted and network-mediated environments. Furthermore, a network-based environment encourages students to search and organize materials related to the topic of study. It also provides a natural collaborative situation for the students to cultivate their social and communicational skills, as many have mentioned. In addition, computers integrate various media and interactive equipment, they can be used to effectively support appropriate tools to guide the process of cooperation among a group of students, to take down the records of the cooperation process, and to analyze the recorded patterns by using artificial intelligence mechanisms.

In this paper, we propose a new learning system, called Learning by Judging. It is a distance learning environment which provides the basic functionality to realize several important concepts in distributed constructionism (Resnick, 1996), such as active learning, consensus development, and meta-cognition. We also integrate the procedure employed in the Delphi method for group decision making, taking advantage of the interactive properties provided by the network environment.

We construct the Learning-by-Judging environment on the WWW, because it provides a multimedia, multi-user environment in which participants from different cultures and professional fields can work together and share their designs and values. In this environment, we expect the users to start participating with minimal amount of prior knowledge so that active learning is encouraged. Then, through the process of construction and re-construction, they will be able to develop general principles and theories.

This approach is an enhancement to project-based learning (PBL) by providing a supplement of peer evaluation. The PBL strategy uses a hands-on project as the focus of learning, so that the students feel what they have learned is integrated knowledge, and it makes sense to them. We believe that the proposed Learning-by-Judging environment has a positive effect in this direction of educational research and development.

II. Research Methodology

This study is based on the CORAL system (Sun & Chou, 1996), which stands for Cooperative Remotely Accessible Learning. We explore pioneering educational theories and new instructional technologies in this study. The research steps are listed in the following.

1. We employ important concepts in constructivism (Jonassen, 1994), collaborative learning (Hooper, 1992; Johnson & Johnson, 1990), and the Delphi method. We consider the characteristics and limitations of network-based computer-assisted learning before we put these concepts together and design the framework of the Learning-by-Judging environment.
2. Based on the theoretic framework, we design a WWW-based Learning-by-Judging interface

and use the Java language for implementation.

3. During Learning by Judging, the learning system automatically records learning behavior in log files. The profile data are then analyzed to generate feedback information for the students and to provide improvement suggestions for the teachers.
4. The students evaluate samples during the learning process. There are two sources for the samples: they can be generated by the students or produced by the system based on domain-specific design principles.
5. Next we conduct instructional experiments on college students in different fields. We investigate the proposed learning environment based on two types of information: (1) log files recorded during learning processes, (2) observations on learner behavior and interviews with the learners.
6. Based on the experiment results, we provide assessment and suggestion for the proposed learning environment and strategy.

III. Design of the Learning Process

The two primary activities in Learning by Judging are sample design and peer evaluation, as described below. For a chosen learning topic, the Learning-by-Judging pedagogy supports a learning procedure of several steps to achieve various learning goals:

1. In the provided Learning-by-Judging environment, a student produces one or more design samples, the samples are then posted via network to all participants. For example, each student produces a blueprint draft on an engineering design course.

2. After all the samples are posted on the network, students start to browse others' products and give a score on each piece of work. The aggregated score of each piece of the work are calculated and posted along with corresponding comments.

3. The students observe the overall scores and comments.

4. Based on the observation of the previous round, each student modifies her or his old samples or re-designs new samples, then repeats the above steps until no further apparent improvement can be achieved or time is up.

5. After the iterative learning cycles are finished, the Learning-by-Judging system analyzes the information collected during the learning process. The analyzed results are informed to individual students as learning feedback for the purpose of meta-cognition.

The Delphi method is frequently employed for group negotiation and decision, we adopted several of its features in the Learning-by-Judging pedagogy. Among them are the anonymous participation of group members, the non-verbal style of message passing, the interactive approach to express suggestive and evaluative opinions across iterative cycles, and finally the accumulative manner of consensus formation. The motivations behind these important features include avoiding the dominance from members with better verbal skills or higher social ranks, alleviating the burden of expressing one's judgment and opinions, learning to appreciate others' work and cooperative techniques, and better knowing the need of time and patience in the consensus-formation process.

By repeating the procedure mentioned above, a student is able to observe the dynamic formation and changes of social consensus through indirect interactions. General design rules are expected to be derived and reported by the student during this process. Since the learning environment records every step of design and scoring, the learning process can be effectively analyzed. For example, the designs and preference of a student at different stages can be categorized. Meta-cognition can be triggered when the student is given this piece of feedback information together with the teacher's comments. The importance of this factor has long been discussed in literature such as (Posner, Strike, Hewson & Gertzog, 1982), we tried in this project to realize it in a distance learning environment.

Also indicated in previous research is this guideline: If only cooperation channels are supported, but no incentives to cooperate are provided, cooperative learning cannot be successfully accomplished (Richartz & Rudebush, 1990; Steeples, 1993). In this aspect, the Learning-by-Judging interface can be viewed as a cooperation-embedded environment because a student starts the sample-generation and peer-evaluation process immediately after she or he enters the environment. In other words, the students start cooperation with each other via network in a natural manner.

IV. Design of Instructional Experiment

In this exploratory study, we focused on the impact imposed on a single student from social interactions. We employed the method of case study to observe the interactions between students, and extracted patterns from the logging data. We analyzed the behavior from three angles: the coherence of the society, the generation of repeated samples, and the movement of preference.

After the implementation of the Learning-by-Judging environment, we designed and conducted two sets of experiments. We had six undergraduate students with different backgrounds to participate in the experiments. We selected the color-matching of clothing as the design task. The two experiments are described below:

1. Fixed number of samples in each cycle: We had three participants (A, B, and C) in this case. Each generated five samples at each of five learning rounds. Besides, each participant evaluated 18 samples during each round, 15 produced by the three participants and three generated by the system according to textbook color-matching principles.
2. Decreased number of samples across cycles: Three subjects (D, E, and F) participated in this experiment. During five rounds of learning, they produced fewer and fewer new samples. They also evaluated 18 samples during each round, some of them were better samples from previous rounds.

As mentioned above, the system recorded each sample generated and each score given by the participants. In the next section we describe the analytical results attracted from the logging data, list some important phenomena observed during the experiments, and summarize the interviews conducted after the experiments.

V. Analysis of Learning Behavior

Coherence with Society

The concept of “knowledge as construction” is emphasized in constructivism-based learning. Based on one’s prior knowledge, a student absorbs new elements during a learning activity to form new knowledge structures. This type of personal mental activity is called “weak constructivism” because the usefulness of the acquired knowledge is yet to be proved. On the other hand, “strong constructivism” emphasizes the acquisition of knowledge from the consensus of a group, i.e., the learning society. This is because interactive experience justifies the usefulness of the knowledge, so it tends to be held stronger in one’s memory. In other words, knowledge construction is not only a personal cognitive process but also a group cognitive process.

Therefore, we proposed two indicators to measure one’s degree of coherence with the learning society from a single student’s point of view. During each learning round, we first cluster the color-matching samples according to their scores and identify a collection of samples with high ranks judged by the group. Similarly, we identify a preference set for each individual. We then compare the preferred samples of a single student and the preferred samples of the whole group to check the overlaps and changes of overlaps between the two sets.

We define the degree of conformity (C) between the "Individual Preference Set (IPS)" and the "Group Preference Set (GPS)" by the following formula:

$$C = \frac{\#(IPS \cap GPS)}{\#(GPS)},$$

in which *#(s)* denotes the size of the set *s*. According to this formula, the degree of one's conformity to the group preference can be interpreted by the value of *C*. More specifically, when *C* equals 1, total agreement between the individual and the group is found; on the other extreme, when *C* equals 0, no similarity between them.

Next, we define the degree of acceptance (A) of an individual by the society as follows:

$$A = \frac{\#(IPS \cap GPS)}{\#(IPS)}.$$

When *A* equals 1, the work pieces chosen by the individual are also fully accepted by the group; on the other hand, when *A* equals 0, none of her or his selections is appreciated highly from the group's viewpoint.

Compared with *A*, *C* is conceptually more suitable to measure how strong the society's opinions affect individual's preference. Thus in the following we will employ *C* as the primary indicator for measuring the coherence of an individual to the learning society and use *A* as an assistant indicator. Based on all logging files, we identified two types of changes in coherence.

1. Approaching coherence gradually: The behavior of four subjects, two from each group, can be categorized into this class. Take student F as an example. In Figure 1 we can observe the conformity value C increases with evolving learning cycles and reaches the peak value 1 in the fourth cycle. This phenomenon clearly reflects the influence from the society. Furthermore, student F's degree of acceptance also climbs up along the learning cycles. This indirectly indicates that the individual preference set moves gradually toward the group set instead of merely enlarging itself to cover the group set.

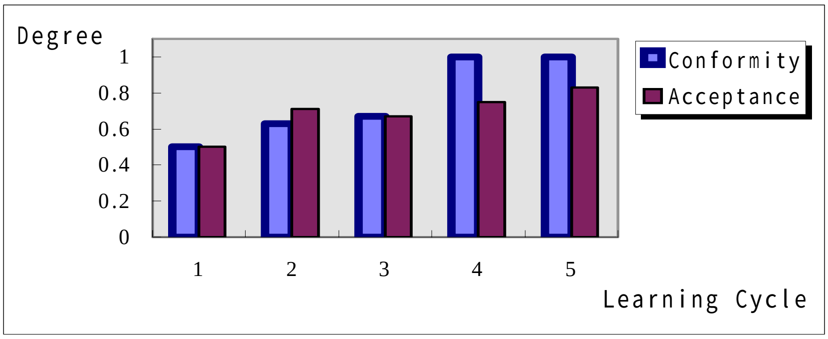

Figure 1. Student F: Conformity and Acceptance Curves.

Figure 2 shows the degrees by student A. The conformity degree also increases. However, compared to student F, student A's degree of acceptance remains low across cycles. A possible explanation is that this student enlarges the scope of preferred samples with new styles learned from the group.

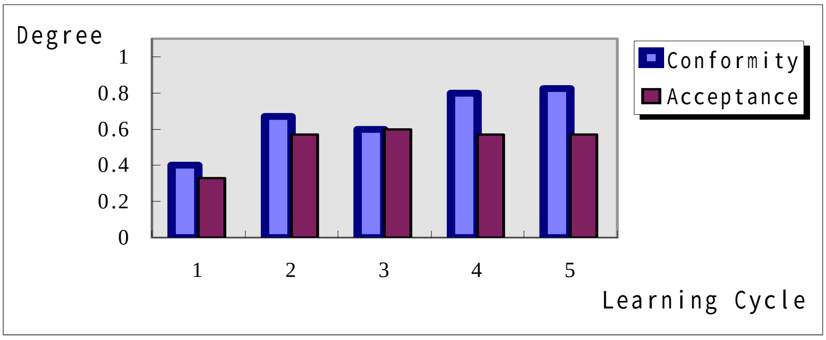

Figure 2. Student A: Conformity and Acceptance Curves.

2. No coherence: Figure 3 shows an example. Student B demonstrates little similarity or overlap with the group. This phenomenon may result from the persistence of existent preference. We will use next analytical item to investigate it in more detail.

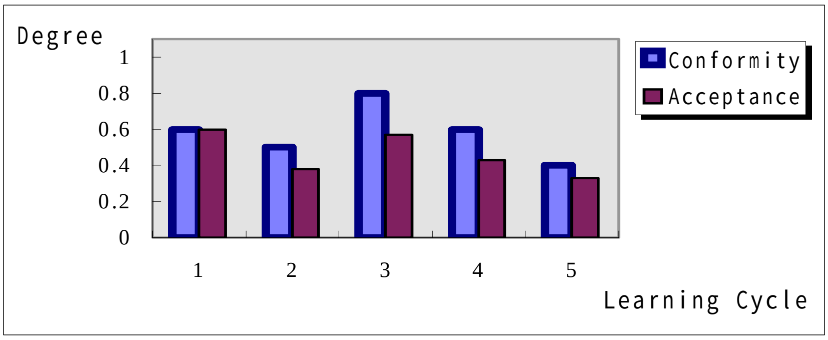

Figure 3. Student B: Conformity and Acceptance Curves.

Repetitive Samples

We observed that some students produced same color-matching samples across cycles. The motivation behind this behavior and its influence on social interaction are interesting research topics. In this study we explore in this direction.

According to the logging data, three students (C, D, and E) repeatedly yielded same samples. They can be further divided into two categories based on the comparison between the grade granted by a student for her or his own work and the average score received from the society.

1. Group score worse than self-evaluation: Students C and D fall in this category. Table 1 illustrates the scoring information of student C. He produced the same pattern in consecutive cycles, and there was a big gap between his own score and the group average. From the interview conducted after the experiment we learned that he had doubts on the group score so he gave the sample another chance to be re-evaluated. However, as we can tell from the table, he gave the repeated sample a much lower score than before. This is another way to show how an individual is influenced by the group.

Table 1. Student C: Repeated Sample

Self Score	Self Rank	Group Score	Group Rank
9	2	5.4	10

6	9	5.4	13

2. Both scores good: Student E produced the same pattern in three consecutive cycles. From Table 2 we observe that both he and the group evaluated highly on this sample. From the interview we knew that he was very confident on his capability of color-matching, and adored that pattern very much. Reinforced by the appreciation from the society, he demonstrated the design consecutively as a result.

Table 2. Student E: Repeated Sample

Self Score	Self Rank	Group Score	Group Rank
10	1	9	1
10	1	9.4	1
10	1	9	1

Change of Preference

Next we investigated how preference changes by observing the preferred samples across cycles. Here we identified apparent difference between the two experiment groups.

We took the top three samples in each cycle, based on the average score, and checked if they re-appeared in the next cycle. We found that in the second experiment, with decreased number of samples across stages, most highly-ranked samples remained in the preferred set across several cycles. In other words, in this experiment group preference is early converged to a small set. On the contrary, in the first experiment, with fixed number of samples at each stage, most top-ranked samples could not survive for more than two cycles. The difference might result from the fact that when students have more opportunities to explore new patterns, they tend to expand their scopes of appreciation.

VI. Observations and Interviews

Based on the observations during the experiments and the interview conducted afterwards, we point out the following items that deserve further investigation.

1. A student indicated that he realized during the learning process that he did not prefer any pattern with black bottom part. Another student found that his evaluative scores were getting closer to that of the group when the experiment proceeded. Because the learning environment supports design-pattern demonstration and peer evaluation, a student can retrospect behavior and observe some changes. Identifying one's own thinking framework via retrospection is a very important feature in constructive learning. We believe that the proposed environment has positive impact in this aspect.
2. Since we conducted the experiments in a synchronous manner, i.e., all participants were on-line and did sample-generation/evaluation in the same time period, some students had to wait for others finishing their jobs. Some of them thus felt impatient with unexpected delay. We suggest two remedies for this issue. One is to have the participant exchange their comments on others' work in addition to evaluation. The other is to conduct the Learning-by-Judging procedure in an asynchronous way, i.e., setting a deadline for each stage so that the participants can switch to other things after finishing their own assignment.
3. Some participants indicated that evaluating too many samples made them nervous. In the mean time, however, some participant felt that if the experiments were cut short they would not be able to identify apparent changes. Again, we found asynchronous

approach may be a good solution to this anxiety problem. We also suggest that more sophisticated designs may prove more suitable in this kind of environment. When there are more things and more complicated situations to consider, a user is likely able to avoid the boredom resulted from giving scores only.

VII. Concluding Remarks

We believe that the concept of learning by judging is worthwhile for further exploration because it shows the following characteristics which provide very good assistance to conventional pedagogy.

1. Learning-by-Judging encourages active learning. Judging is a relatively easy activity that can alleviate pressure and encourage early involvement. Giving scores is also simple in terms of expression, thus expressive ability is no longer such a crucial factor as observed in peer competition in traditional classrooms. Furthermore, anonymity provides a comfortable environment for trial-and-error.
2. Learning-by-Judging promotes information sharing and social interaction. Peer evaluation is a cooperation-embedded activity. Participants share experiences and opinions via network in a natural way.
3. Learning-by-Judging helps acquiring integrated knowledge. Both design and evaluation are the results of high-level knowledge integration. Moreover, participants can identify their existent cognitive schemes through observation and comparison. Meta-cognition and meta-learning can be achieved in this kind of learning environment.
4. Learning-by-Judging benefits instruction management. Based on the logging files provided by the system, a teacher can check the progress of individual students. When a student’s cognitive or communicative patterns are identified, necessary aids can be provided by the teacher in time.
5. Learning-by-Judging supports feedback information. By using indicators such as those introduced above, the system is able to give students useful information about their preference, their ways of interaction, and the influence from the society. Relevant design principles can be explicitly explained after their preferred samples are analyzed. Knowledge acquired this way makes sense to the students and is proved to be more useful than simple drill and practice.

Finally, we suggest that the Learning-by-Judging environment can be improved along the following dimensions.

1. To support both synchronous and asynchronous learning.
2. To integrate various communication channels so that more collaborative information can be exchanged during learning.
3. To connect the environment to relevant hypermedia courseware so that participant can get appropriate assistance.

Reference

Hooper, S., (1992). Cooperative learning and computer-based instruction. Educational Technology Research & Development, 40(3): 21-38.

Johnson, D. W., Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (ED.) Cooperative learning: theory an research, New York: Praeger Publishers, 23-37.

Jonassen, D. H. (1994). Thinking technology: Toward a constructivism design model. Educational Technology, 34(4), 34-37.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accoummodation of a scientific conception: toward a theory of conceptual change. Science Education, 66, 211-227.

Resnick, M. (1996). Distributed Constructionism. Proceedings of the International Conference on the Learning Sciences.

Richartz, M. & Rudebush, T. D. (1990). Collaboration in hypermedia environments. In D. H. Jonassen & H. Mandl (Eds.), Design hypermedia for learning, Berlin: Springer-Verlag, 311-317.

Steeples, C. (1993). A computer-mediated learning environment for adult learners: Supporting collaboration and self-direction. Journal of Educational Multimedia and Hypermedia, 2(4), 443-454.

Sun, C.T. and Chou, C. (1996). Experiencing CORAL: Design and Implementation of Distant Cooperative Learning. IEEE Transactions on Education, 39(3), 357-366.