Development and Evaluation of PDF Report Annotation Tool GABA Facilitating Comment Reuse

Tetsuro Kakeshita
Department of Information Science
Saga University, Saga, 840-8502, Japan

Shoichi Motoyama
NEC Corporation, Tokyo, Japan

ABSTRACT

Comparing online and paper-based environment for report submission and correction, the former supersedes to the latter, since (1) the turn-around time becomes shorter, (2) teaching opportunity increases, and (3) as a consequence, the student's achievement level becomes higher in the online environment.   In this paper, we propose an annotation tool GABA for PDF document in order to reduce correction time by the teachers and to facilitate instruction to students.   In a usual class, the same or similar assignments are given to the students.   Then it is often the case that many students make similar mistakes.   A teacher can register and classify common correction comments to GABA.   Report correction time becomes significantly shorter by reusing the registered comments. GABA also provides various support functions in order to assist efficient checking of numerous report files such as (1) sorting of frequently-used comments, (2) similarity-based file sorting, and (3) cross tabulation of comments using category and weight.

Key words*: e-Learning, report correction, annotation, PDF document, reuse of comments*

1. INTRODUCTION

It is important to give assignments to students and to review and correct reports submitted by the students at a college level education in order that the students carry out in practice what they have learned.   Student's achievement level will become higher, as the feedback from the teacher is more concrete and detailed.   However such reviewing and correction of the submitted reports requires much time and effort of the teacher so that a student may not have enough feedback if the teacher has many students registered to the current course.

As e-learning become popular, more assignments are given through an online environment and more reports are submitted electronically.   Comparing an online report submission and feedback environment with traditional offline environment, the online environment is better than the offline one since (1) the turn-around time becomes shorter from the student's viewpoint, (2) the opportunity increases to teach students, and (3) as a consequence, the student's achievement level becomes higher in the online environment [1].

In this paper, we develop an annotation tool GABA for PDF document and combine it with Moodle assignment module in

order to reduce correction time by the teachers and to facilitate instruction to students.   It is easy to generate PDF file from various types of reports such as Latex/Word document, PowerPoint presentation, C++/Java program and UML diagrams.   Thus GABA can be used to a wide variety of reports.   In a usual class, the same or similar assignments are given to the students.   Then it is often the case that many students make similar mistakes.   A teacher can register and classify common correction comments to GABA.   Report correction time becomes significantly shorter by reusing comments.   The list of registered comments is a set of checkpoints for report evaluation so that the fairness of evaluation is achieved simultaneously.

Adobe Acrobat Professional provides the annotation/review tool in order to add various types of annotations (note, text, etc.) to arbitrary places of a PDF document.   However the user needs to type the comment text using keyboard.   GABA extends this function by utilizing Acrobat SDK [2].   In addition to the comment registration and insertion functions, GABA provides the following support functions to facilitate review, correction and evaluation of the reports.

1.   Classification of the registered comments.
2.   Sorting of comments using frequency in order to speed up the insertion of the frequently used comments.
3.   Retrieval of the similar reports of the current report in order to reuse the annotations of the current report.
4.   Cross tabulation of the used comments according to the

category and the importance level of the comment. This function is used to evaluate the current report and to analyze the achievement of the entire class.

GABA has advantages over related annotation tools. Conventional report annotation system such as CoCoA [3] requires installing custom software to the teacher's and student's computers. It is often the case that the target document format is restricted. Although it is known that annotation is useful for knowledge sharing [4], many educational applications focus on annotation to the teaching materials such as lecture note, handouts and slides.

This paper is organized as follows. We introduce GABA and explain GABA annotation functions in Section 2. In Section 3, we discuss the evaluation of GABA annotation function. Other support functions of GABA are explained in Section 4. Section 5 introduces the report submission and feedback system using Moodle assignment module. The system is used to efficiently manage submission, correction and feedback of the reports.

2. GABA ANNOTATION FUNCTION

We developed GABA as a plug-in of Adobe Acrobat Professional. GABA is installed at the plug-in folder of Acrobat. GABA is developed using Acrobat SDK and C++ programming language running under Microsoft Visual Studio 2010. Acrobat SDK provides various APIs to customize menu, to access PDF document, and to handle events [2].

GABA has two types of functions: annotation function and support functions. We shall introduce GABA and explain the annotation function in this section.

GABA can be used to put comments to various PDF documents. GABA maintains correction comments classified by categories. A user, usually a teacher, first adds empty comments to a PDF document using the annotation function of Acrobat Professional. The user then selects an annotation to choose a category and an associated comment from the GABA menu (Fig. 1). The GABA menu contains submenus corresponding to the categories. The category submenu contains comments belonging to the category. GABA inserts the selected comment to the annotation. A user can add any number of comments to an annotation.

The user can register a new category or comment to the GABA menu by selecting the corresponding menu item. The category name is specified at a dialog box. A comment is first added to an annotation and then can be registered to an arbitrary submenu. The user can associate an integer (weight) to each comment in order to describe the importance of the comment.

When a user opens a PDF file, GABA reads the correction comment file and category file to build the GABA menu. The correction comment file is a CSV file containing comment title, comment text, category id and the weight of each comment. The category file is a CSV file containing category id and category name. The two files are stored at the same folder as that of the PDF file so that the teacher can easily switch the set of correction comments and categories for each folder by creating folder corresponding to each assignment.
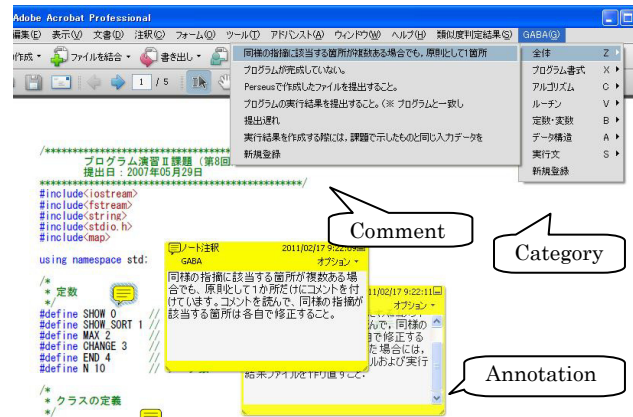

Fig. 1 GABA User Interface

Table 1. Example of GABA Categories and Comments

Category	# of Comments
Miscellaneous	6
Coding Format	9
Algorithm	10
Routine	2
Constant and Variable	10
Data Structure	5
Statement	6

Since the comment and category files are CSV files, they can be easily created and edited using spread sheet software such as Microsoft Excel. It is desirable to predefine categories and comments before the report evaluation. One reason of this is to assure fairness of report evaluation by the teacher and the teaching assistants (TA). The other reason is that better report will be produced by the students by disclosing the categories and comments as evaluation criteria. We illustrate predefined categories and the comments in Table 1. These are used to evaluate C++ program as illustrated in Fig. 1. These comments are defined based on our guidelines for C++ programming [5].

In order to speed up the operation to add comments to an annotation, keyboard shortcuts are assigned to each category. The shortcut keys are selected within the area of the left hand. The user inserts comments using pointing device. A user can select an annotation by the right hand (using mouse) and can insert a comment string using the left hand. Such separation is useful to speed up the comment addition operation.

GABA also provides right-click menu which a user can quickly select GABA menu. The right-click menu appears when the user selects an annotation. The selected comments are inserted to the end of the annotation so that the movement of the mouse cursor is minimized.

3. EVALUATION OF GABA ANNOTATION FUNCTION

We compare the operation time for comment addition using Acrobat 8 and GABA. We assign 10 PDF files to 9 examinees. The PDF files are C++ programs of approximately 350 lines (5 pages). There are several degrees of completeness among the programs. Among the examinees,
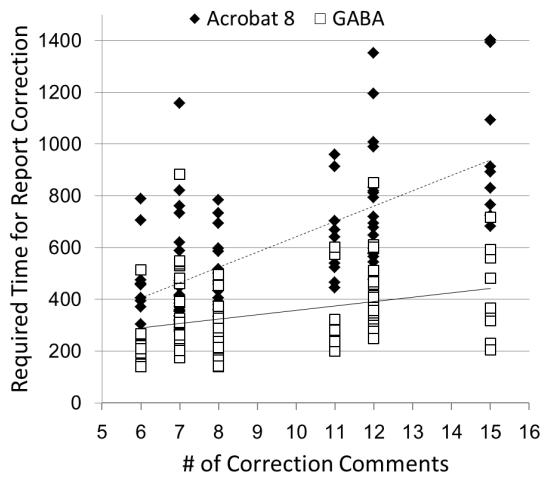
Fig. 2 Operation Time of Acrobat 8 and GABA for Comment Addition

6 are professors and 3 are graduate students having experience as TA. Since we intend to evaluate the operation time and do not intend to measure the time to consider comment text, we give instructions to the examinees about comments and their places for each program. Fig. 2 illustrates the required time and the number of correction comments for each pair of examinee and program.

It can be observed that the required time grows according to the growth of the number of comments. In case of Acrobat 8, the correction time grows 52.2 seconds for each comment. On the other hand, the correction time grows only 11.5 seconds for each comment by using GABA. The standard deviation of the time is smaller for the case of GABA. This is mainly because that the examinees do not require typing of comment text when the comment is registered to the system.

We also collect opinions of the examinee of the above experiment. All examinees agree with the effectiveness of GABA. They also admit the convenience of the right-click menu. We also have the following comments from the examinees. The GABA support functions are designed based on the comments.

- It is desirable to provide the sorting function of the comments by the usage frequency.
- It takes time to understand correspondence between category and comment. Comment addition becomes faster after the examinee understands the correspondence.
- It is desirable to show a list of candidate comments when the examinee selects some text of the PDF file.

4. GABA SUPPORT FUNCTIONS

Considering the comments collected from the examinees, we improve GABA usability by developing various support functions. We shall explain such functions in this section.

Sorting of Frequently-used Comments: GABA maintains a LRU (least recently used) queue of the comment usage and shows the list of frequently used comments at the GABA menu (Fig. 3). Frequently used comments are also shown at the

right-click menu. This function is useful to quickly add frequently used comments. Comment addition time from the quick access menu is 45% of the time from the regular GABA menu according to our evaluation result.
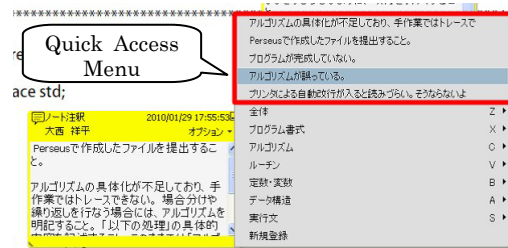

Fig. 3 Quick Access Menu for Frequently Used Comments

Similarity-based File Sorting: GABA computes similarity of the PDF files stored in the current folder. The user can view a sorted list of files based on the similarity to the current file. We have observed that similarity of the comments associated to a file increases according to the increase of the file similarity. Combination of the sorted list of similar files and the quick access menu for frequently used comments is useful to maximize comment reuse and to reduce comment addition cost.
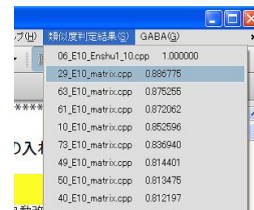

Fig. 4 Similarity-Based File Sorting

The similarity score between files is calculated as follows. At first, we summarize the characteristics of each file using a hash histogram. Considering a file as a string, a hash value is calculated for each fixed length substring starting from each character of the file. The hash histogram is defined by the number of substrings corresponding to each hash value. The similarity score is then calculated by the correlation coefficient between the two hash histograms.

Cross Tabulation of Comments using Category and Weight: Each registered GABA comment is assigned an integer weight. GABA provides cross tabulation of the number of comments using category and weight (Fig. 5). Such cross tabulation is useful to evaluate the current PDF file. GABA also provides cross tabulation of the comments added to the files of the current folder. Such cross tabulation is useful to understand the degree of achievement throughout the class.
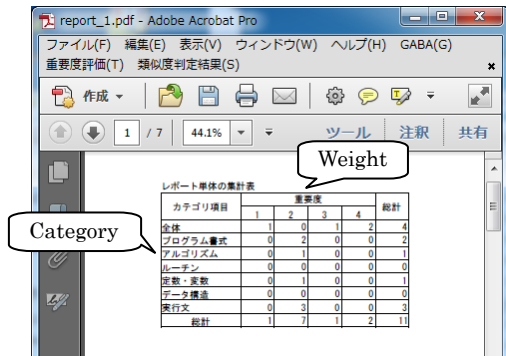
Fig. 5 Cross Tabulation of Comments

5. ONLINE REPORT SUBMISSION SYSTEM

An online report submission system is necessary to fully utilize GABA.   This is because (1) students can submit reports at any time and place, (2) the teacher's load is minimized to receive the reports and to return the corrected reports, (3) students can receive the corrected report as soon as possible, and (4) the received reports and the corrected reports are stored for future reference.

We use Moodle assignment module to satisfy these requirements.   The overall structure of the online report submission system is illustrated in Fig. 6.
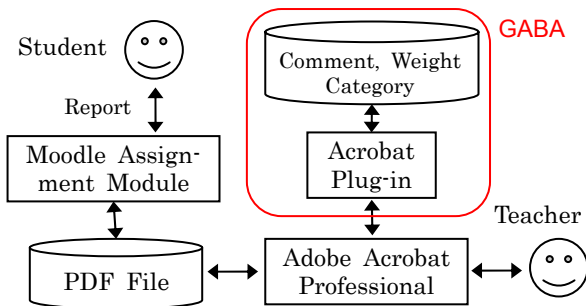

Fig. 6 Overall Structure of the Online Report
Submission System

The teacher can add an assignment at any week of the Moodle course page.   Each assignment page contains assignment name, assignment description (e.g. instruction to students), grading scale, due date, assignment type, etc.   If the teacher selects "advanced uploading of files" assignment type, it becomes possible to return a response file to each student.

Each student uploads the report file (PDF format) from the assignment page.   Since each student registers to the Moodle course at the beginning of the semester, it is possible to manage each student separately.

The teacher can download the submitted report files at the teacher page of the assignment.   The page contains the information of the submission date and time of each report so that the teacher can easily know the students who do not submitted the report or who missed the submission deadline.
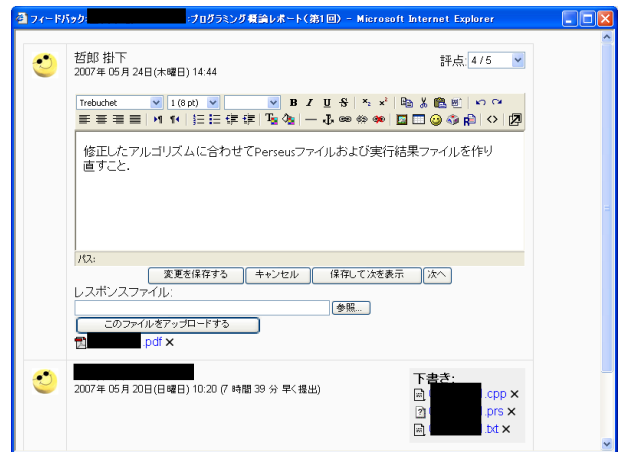

Fig. 7 Sending Score and Feedback File to Students

GABA is then used to check the downloaded reports.   The corrected reports are returned to each student as the response file through the assignment page (Fig. 7).   The evaluation result is notified to each student at the same time.   An online report submission system allows both of the teacher and students to have a copy of the report files before and after correction.   This is useful to systematically accumulate evidence for accreditation such as ABET (USA) or JABEE (Japan).

6. CONCLUSION

We are using GABA in combination of Moodle.   Students have their own note PC and submit PDF reports using Moodle assignment module.   It is easy to create PDF report from various types of applications so that we use GABA for lectures other than computer programming.   Our Moodle site can be accessed from the internet.   Teachers check the reports using GABA and return the annotated reports using Moodle.   The annotated reports are also used as the evidence of the evaluation and accreditation of our education program.   In this sense, combination of GABA and Moodle is a powerful tool for quality assurance of higher education.

Future topic on this research includes automatic checking and annotation of the submitted reports.

REFERENCES

[1]   N. Sumi, X. Li, H. Deguchi, T. Ohta, S. Sakai, "Analysis of turnaround time in on-line report correction support system",   Journal of Information Processing Society of Japan, 48 (8), 2007, pp. 2781-2790.

[2]   *Adobe   Acrobat   SDK,* http://www.adobe.com/devnet/acrobat.html, 2012.

[3]   Y. Yano, H. Ogata, R. Sakakibara, R. Wakita, "CoCoA : Communicative correction assisting system for learning Japanese compositions",   Transactions of Japanese Society for Information and Systems in Education, 14(3), 1997, pp. 21-28.

[4] S. Ito, Y. Sumi, K. Mase, S. Kunifuji, "SmartCourier: An annotation system for adaptive information sharing", Journal of the Japanese Society for Artificial Intelligence, 17(3), 2002, pp. 301-312.

[5] T. Kakeshita, "*Guidelines for C++ programming*", http://www.cs.is.saga-u.ac.jp/syllabus/ GuideLine/Cguide.html

Tetsuro Kakeshita
He received the B.S., M.S and Ph.D. in computer science from Kyushu University, Japan in 1984, 1986 and 1989 respectively. Since then, he has been with Saga University. His main research interests include complexity analysis of database and software, and quantitative analysis of ICT education.

Shoichi Motoyamaa
He received the B.S. and M.S. in computer science from Saga University, Japan in 2007 and 2009 respectively. Currently he is working for NEC Corporation as an IT engineer.