CHAPTER 12

Learning Support Systems Based on Cohesive Learning Analytics

Fumiya Okubo, Masanori Yamada, Misato Oi, Atsushi Shimada, Yuta Taniguchi and Shin’ichi Konomi

1 Introduction

E-learning systems have become indispensable in educational institutions (Alsabawy, Cater-Steel, & Soar, 2013). The Learning Management System (LMS), the Course Management System (CMS), or the Virtual Learning Environment (VLE), is one such e-learning system and it is widely accepted in higher education (Islam, 2015). LMSs are used for facilitating e-learning. The LMS process is used to store and disseminate educational material and to support administration and communication associated with teaching and learning (McGill & Klobas, 2009). Ubiquitous technology is effective in enhancing the effects of the LMS and VLE on learning performance (Yin et al., 2015; Yamada et al., 2016). By using logs stored in the server, information technologies allow instructors to understand student learning behaviors. By using various data from sources such as the logs, learning analytics research contributes to the clarification and improvement of education and the learning environment (e.g., Ifenthalar, 2015; Ogata et al., 2015).

To support student learning, revealing its patterns by using learning analytics methods might be useful. For example, Oi and her colleagues (Oi et al., 2015; Oi, Okubo, Shimada, Yin, & Ogata, 2015; Oi, Yamada, Okubo, Shimada, & Ogata, 2017) investigated the relationship between the use of e-books outside the classroom and academic achievement. They categorized e-book logs from the e-book delivery system as follows: If a log was recorded before a class session in which the same e-book was used as the textbook, it was labeled a preview log, and if after, a review log.

As pointed out by Ausubel (1960) in his classical study, if a student gains knowledge of the contents of a course by performing a preview, it may help the student’s learning in the classroom by acting as an advance organizer. Performing a review may facilitate memory consolidation by working as a rehearsal of the contents of the course. The main findings are as follows: (i) The preview is more deeply related with academic achievement than the review (Oi, Okubo
et al., 2015), (2) high achievers may actively link related e-books and their pages with a preview and a review (Oi, Yin et al., 2015), and (3) relatively low achievers attempted to perform previews, but they tended to give up easily (Oi et al., 2017).

These results suggest that previews may be effective for higher academic achievement, and in indicating some issues, such as, revealing why low achievers gave up on their review and how to help them start their preview and review processes. In Sections 3.2 and 3.4, we will introduce possible solutions.

The e-book log also provides useful information by which to predict a learner's psychological state such. One common issue under discussion is how psychological variables affect learning performance within the learning environment using ICT (Greene & Azevedo, 2010). Psychometric data as well as learning logs should be collected, in order to analyze learners' behaviors in order to achieve effective learning support. Particularly helpful would be introducing learning styles, such as self-regulated learning (SRL) (Roll & Winne, 2015).

Yamada et al. (2015) indicate that self-efficacy, one element of SRL, had a significant correlation with learning behaviors, such as highlighting and annotation. Yamada et al. (2017b) further suggested that the use of cognitive learning strategies, such as annotation, as well as giving appropriate time for reading learning materials, play important roles in enhancing SRL awareness. Yamada et al. (2017a) also indicated that awareness of self-efficacy, intrinsic value, and cognitive learning strategies had a significant correlation with the frequency of out-of-class activities, submission times of reports, and learning performance. Moreover, they found that the relationship between SRL awareness and out-of-class activities such as, reading activities and bookmarking, had a significant correlation with SRL awareness.

Due to the high adoption of LMS that promote the massive store of learning data, as well as the analyses of this data, the findings mentioned above could bring us insights into student activity and indicators of teaching quality. In the following section, we introduce an integrated ubiquitous learning platform, “M2B (Mitsuba),” in order to improve learning and teaching based on learning analytics.

2 M2B: Learning Support and an Analytic Platform

M2B consists of a common learning management system (Moodle), an e-portfolio system (Mahara), and an e-book system (BookRoll) (Ogata et al., 2017). Moodle is mainly used by teachers to manage student attendance, provide quizzes, and
receive reports. Both teachers and students use Mahara for keeping notes in e-portfolios. BookRoll enables students to browse e-book materials before, during, and after lectures, anywhere and anytime, using their PC or smartphones. All user actions performed on the e-book system, such as page flips and opening material, are recorded in the learning logs and automatically sent to the university’s database. The e-book system provides additional functions of bookmarking, highlighting, annotating, and searching. By the end of April, 2017, approximately 45,000,000 logs were collected from approximately 20,000 students. Analyses of educational big data from M2B showed progress in the students’ understanding of activities. In the next section, we introduce M2B plug-ins based on the learning analytics methods used to enhance learning performance.

3 Learning Support Systems Based on Cohesive Learning Analytics

3.1 A Visualization Tool to Overview Learning Activities through a Course

Many kinds of logs of students’ learning activities are stored in the M2B system. For teachers to utilize these learning logs and improve their teaching in their classes, specific purposes should be used to visualize the logs. Here, we consider a system of visualization in order to understand an overview of student learning activities during a course.

At Kyushu University, the training of students and the evaluation of student learning both inside and outside the classroom are regarded as important issues. In order to analyze and visualize the many kinds of logs from this point of view, we select nine major learning activities, and evaluate them by individual students, using a scale of 0 to 5 points for each week of the course. A vector of these nine evaluations is called an Active Learner Point (ALP). The nine selected learning activities and the method for evaluating them are summarized in Table 12.1. We note that

– The logs of attendance, quizzes, reports, and course views are stored in Moodle.
– The logs of slide views, markers, memos and actions are stored in BookRoll. Here, the item “slide views” is calculated by the total time of slide views outside of class.
– The logs of word counts in journal are stored in Mahara.

The logs of the number of course views, total time of slide views, number of markers, number of actions, and word count in a journal are evaluated relatively in a course. For example, the top 10% of students who are most active in their learning activities obtain a score of 5.
TABLE 12.1 Criteria for active learner point (ALP)

<table>
<thead>
<tr>
<th>Activities</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance</td>
<td>Attendance</td>
<td>Being late</td>
<td>absence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quiz</td>
<td>Above 80%</td>
<td>Above 60%</td>
<td>Above 40%</td>
<td>Above 20%</td>
<td>Above 10%</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Report</td>
<td>Submission</td>
<td>Late submission</td>
<td>No submission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course Views</td>
<td>Upper 10%</td>
<td>Upper 20%</td>
<td>Upper 30%</td>
<td>Upper 40%</td>
<td>Upper 50%</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Slide Views</td>
<td>Upper 10%</td>
<td>Upper 20%</td>
<td>Upper 30%</td>
<td>Upper 40%</td>
<td>Upper 50%</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Markers</td>
<td>Upper 10%</td>
<td>Upper 20%</td>
<td>Upper 30%</td>
<td>Upper 40%</td>
<td>Upper 50%</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Memos</td>
<td>Upper 10%</td>
<td>Upper 20%</td>
<td>Upper 30%</td>
<td>Upper 40%</td>
<td>Upper 50%</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Actions</td>
<td>Upper 10%</td>
<td>Upper 20%</td>
<td>Upper 30%</td>
<td>Upper 40%</td>
<td>Upper 50%</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Word Count</td>
<td>Upper 10%</td>
<td>Upper 20%</td>
<td>Upper 30%</td>
<td>Upper 40%</td>
<td>Upper 50%</td>
<td>Otherwise</td>
</tr>
</tbody>
</table>

The ALPs of students represent their activeness for learning in the course. Therefore, the ALP can also be useful for predicting students' performance, an important issue in the research field of educational data mining (Baradwaj & Pal, 2011). Our research group showed that the students' final grades (Okubo et al., 2017a, Okubo et al., 2017b) and quiz scores (Okubo et al., 2018) could be predicted, using the ALP and Recurrent Neural Network.

For visualizing an overview of students' learning activities using ALP, we have developed a system called the Active Learner Dashboard, which consists of three functions. Each function enables teachers and students to know the learning expectations in the course. Here, we explain the implementation of the system and its details. The Active Learner Dashboard is implemented as the set of “block” type plugins in Moodle. The “block” type plugins are always displayed in the sidebar during the course as simple “block display” mode (see Figure 12.1).

The advantage of this display mode is that when teachers and students access the course, they can readily view their teaching and learning situations, respectively, at a glance. By clicking the “View” button, a “detailed” mode of
each function is displayed. Figure 12.2 shows the detailed displays of the Active Learner Ranking for teachers (a) and for students (b).

This function displays the ranking of each student according to the sum of the ALP that they have acquired so far in the course. In order to maintain student privacy, we have different contents for teachers and students. For teachers, the breakdown of points obtained for each item in the ALP is displayed. For students, their own total points and rank are displayed, along with the top three students.

Figure 12.3 gives a detailed display of the Active Learner Process (a) and the teacher’s Active Learner Distribution (b) for teachers. The Active Learner Process shows a transition in the average value of ALPs each week as a red line in the graph. A breakdown of the ALP is displayed by mouse-over. By selecting each learning activity, the rate of achievement is displayed in a bar graph. The Active Learner Distribution shows the distribution of students, categorizing them into specified numbers of levels based on their ALPs for each week. By using these functions, teachers can see how the activity of their class changes week to week.

We asked 77 students attending the “Introduction to Pedagogy” course that opened the first semester in 2016, to use the system of Active Learner Ranking and then we gave them the following questionnaire:

(1) Were you concerned about your ranking in the Active Learner Ranking?
(2) Did you reflect on your learning awareness by using the Active Learner Ranking?
(3) Did you care about the learning activities of other students due to the Active Learner Ranking?
(4) Did competition awareness increase due to the Active Learner Ranking?
(5) Did you actively learn in this course due to the Active Learner Ranking?

The responses are summarized in Figure 12.4. We can see that approximately 60% of the students were concerned about their Active Learner Ranking, they reviewed their learning awareness, and they were conscious of other students’ learning activities due to the system. On the other hand, students who said that the competition awareness increased or that they learned actively in the course was less than 50%.

These results imply that the system of Active Learner Ranking is effective for changing student learning awareness, however it does not reach the same effectiveness in the actual change in student learning activities. For this reason, it seems necessary to consider a system that responds to each student individually according to their situation.
3.2 *Automatic Summarization of Learning Materials for Enhanced Preview*

When discussing enhancements to the learning processes, it is often argued that studying in advance of class is very important to enable students to understand the narrative of the class, become familiar with important keywords, and discover new terms and concepts. Ausubel (1960) emphasized the importance of providing preview information to be studied in advance.

In addition, Beichner (1995) reported that adequate preparation prior to lectures leads to improved student performance. In universities, students are often asked to prepare for their next class by reading a textbook or previewing material. Hereafter, we use the term “preview” to denote any form of studying and/or reading material provided in advance by a teacher.

The proposed automatic summarization system was designed to produce a set of lecture slides that contain text explanations, as well as various types of visual content, including images, tables, and mathematical formulas. The purpose of slide summarization is to maximize the importance of the content using a selection of subset pages under a given condition (in this case, browsing time). The most important pages must be selected without losing the overall narrative of the lecture. In the proposed system, we assume that the narrative of the original lecture material consists of a sequence of important keywords. Moreover, we assume that important pages exhibit the following characteristics:

- Sufficient content to be worth browsing,
- Unique content,
- Keywords that appear frequently in a page,
- Keywords that rarely appear throughout the set of pages.

Based on the first assumption, the system selects pages containing figures and/or tables that support the reader’s understanding of the content. Based on the second assumption, the system removes redundant content, such as animations. Based on the third and fourth assumptions, the system locates important keywords that appear frequently in a given page, but rarely appear throughout the total set of pages. These characteristics are analyzed using a combination of image and text processing.

Figure 12.5 shows an overview of the proposed approach. First, a set of slide pages, $S$, are analyzed to extract important visual and textual features from each page. In terms of visual importance, how much text and how many figures, formulations, or other objects are contained in each slide is estimated using a background subtraction tool, and an inter-frame difference tool. In addition, word importance is measured using the $\text{TF-IDF}$ (term frequency–inverse document frequency) method (Salton, 1988; Wu, 2008). Meanwhile, a
teacher estimates the time that students need to spend studying each slide. Then these visual, textual, and temporal features are combined to predict an importance score $I(S_j)$ for each slide page, where $S_j$ indicates the page number of the set of slide pages, $S$. Finally, an optimal subset $\hat{S}$ is selected, whereby the importance score is maximized for a given preview time. For more details, refer to Shimada (2017).

We investigated the effectiveness of the proposed approach in a series of information science courses over three weeks. In total, 372 first-year students, including both art and science students, attended the classes, which began in April, 2015. All students brought their own laptops to class every week. Each week, students were asked to preview material in preparation for the next lecture. We prepared three sets of summarized slides (short, medium, and long versions), in addition to the complete set of slides. Figure 12.6 displays the full set of slide pages from the third week, along with the summarization results.

The four sets of slides outlined above were given to four student groups using an e-book system (Yin, 2014). This system enables us to collect browsing logs for each set of slide pages to confirm whether a student actually previewed the material within the specified period (the week preceding the lecture). Therefore, we can classify the students into five groups, including those who...
We compared the quiz scores among the five groups:
- Short: students who previewed the short set of summarized slides
- Medium: students who previewed the medium set of summarized slides
- Long: students who previewed the long set of summarized slides
- All: students who previewed the original set of slides
- None: students who did not preview any slides.

We then conducted t-tests to compare all possible pairings of the groups. We compared the groups who previewed the three summarized slide sets with those who previewed all the slides. We also compared the groups who previewed sets of slides with the group who did not preview any slides. Figure 12.7 shows the average score for each group.

The average scores for groups that previewed the various sets of slides (short, medium, long, and all) exceeded those of the group that did not preview the slides (none). With regard to the first week, the contents of the slides were not complex, so there was little difference in the quiz scores. The contents of the slides became more difficult as the lecture series proceeded, and therefore, the differences between the scores increased over the second and third weeks.

The most important point to note (and our major argument), is that the summarized slide sets did not have a negative effect on the quiz scores. This is
supported by the fact that the null hypotheses between the summarized slides and all the slides were not rejected. Instead, previewing the short and medium sets of summarized slides improved the students’ quiz scores compared with previewing the full set of slides (all).

From these results, we can see that the proposed method for summarizing the slides is very effective. In addition, students who previewed the medium set of summarized slides obtained the best scores in all of the quizzes. Note that this summarized slide set was given to a different group each week. This is a very interesting result, and is closely related to the achievement ratios obtained in previewing the slides.

Figure 12.8 shows the achievement ratio for each slide set in each week. With regard to the second week, almost 90% of students previewed their slide sets, regardless of the length of the set. However, in the cases of the first and third weeks, achievement ratios clearly decreased as the number of slide pages and the estimated preview time increased (see Table 12.1). Almost one-half of all students who were given all the slides abandoned their preview early in the process. We concluded that too much material caused a decrease in the students’ attention spans. Therefore, it is important to make a summary set of slides for preview by students.

3.3 Real-Time Analysis to Understand Students’ Behaviors during a Class

There are roughly three types of feedback loops in terms of their frequency: yearly, weekly and real-time. A typical example of a yearly (or term-by-term)
feedback loop, is the assessment and improvement of education. Student grades, examination results, class questionnaires, and so on, are typically analyzed and evaluated (Mouri, 2016; Okubo, 2016). The yearly feedback loop is designed so that the results will be delivered the following year (or term). Students and teachers will not directly receive the feedback results.

A weekly feedback loop can recommend related material based on student status determined using the prediction of academic performance through the analysis of the learning logs, such as attendance reports and quiz results (Shimada, 2016). In contrast to the yearly feedback loop, the analysis results are directly fed back to the students and teachers.

The main difference between a weekly and a real-time feedback loop is that the results can be fed back to the students and teachers in the classroom during a lecture. The teacher can monitor what students are doing (e.g., whether students are following the explanation or whether they are doing something unrelated to the lecture). Based on this timely feedback, a teacher can easily control the speed of the lecture, take more or less time for exercises, or make other adjustments.

Our study has focused on feedback and specifically, how to feedback efficient information to the on-site classrooms during lectures. The aim of this research is to realize real-time feedback, which has not often been discussed with respect to the on-site environment. Our target is on-site classrooms where teachers give lectures and a large group of students are listening to the teacher’s explanation, participating in exercises, etc. In such a large classroom, it is not easy for teachers to grasp each student’s behavior and activities.

We utilize an e-Learning system and an e-Book system to collect real-time data regarding learning activities during the lectures. We have developed two main feedback systems. One is useful for the teacher prior to beginning the
lecture. This system provides summary reports of the previews of given materials and quiz results. Using the preview achievement graph, the teacher is able to check which pages were well previewed and which pages were not. Additionally, the teacher can check which quizzes were difficult for students, and the suggested pages that should be explained in the lecture to aid students.

The other is the real-time analytics graphs, which are helpful for the teacher to control his/her lecture speed. The system collects e-Book logs operated by students sequentially, and performs analytics in real-time on how many students are following the teacher’s explanation.

In Figure 12.9, we present an example case study which was applied to a lecture in our university. The timeline is divided into two parts: before starting a class and during a class. During the previous lecture, the teacher gave students some preview materials that were automatically generated by the summarization technique (Shimada, 2017). Students previewed the given materials before the class and the operation logs recorded during the previews were collected in the system. Before the class started, students answered the quiz questions and the results were collected in the server.

Just before the lecture started, our system analyzed the learning logs to create a summary report containing previews of the achievement and quiz results. Additionally, the system provided information regarding important pages that should be explained well in the lecture. For example, should the teacher focus on pages that are related to the quiz questions, especially those that have led to
lower quiz scores? Our system analyzed in advance the relationship between quiz statements and their related pages in the lecture material.

During the lecture, the teacher explained the content of the materials and students browsed the materials on their lap-tops. In our university, students are asked to open and browse the same page as the teacher, and to highlight or add notes on the important points. During the lecture, learning logs were sequentially collected and stored. The analysis results were immediately visualized on the web interface, as shown in Figure 12.10, and updated every minute. Therefore, the teacher was able to monitor the latest student activity. The visualization included real-time information regarding how many students were following the lecture, how many students were browsing previous pages, etc. The teacher adaptively controlled the speed of the lecture according to what he/she was learning about the students. For example, if many students were not following the lecture and were still on previous page, the teacher slowed down the lecture. For detailed information about implementation, refer to Shimada (2017b).

We investigated the effectiveness of the proposed system in two classes at Kyushu University, Japan. One was the control group (N = 58), without the system. The other was the experimental group (N = 157), using the system. The contents of two lectures were completely the same. The class was designed to provide an introduction to information and communication technology in a number of disciplines. First-year students, including both arts and science students, attended the class, which began in October, 2016. All of the students brought their own laptops to class. The lecture was given by the same teacher, using the same materials.

![Real-time heat map of browsed pages](image-url)
The analyses of the preview status and the quiz scores were performed just before the lecture started. The system reported that most students answered two of eleven questions incorrectly. The pages related to the quizzes were shown on the display and the teacher confirmed them. The teacher spent considerable time on the explanation of these pages. In fact, the page was opened by the teacher for three minutes in the experimental group, meanwhile, one minute for the control group. We analyzed the number of bookmarks, highlights and notes on the two pages, where the teacher emphasized the explanation. In the experimental group, about 61% of students used those functions. On the other hand, only 53% of the students in the control group used the functions.

When the teacher gave the lecture to students in the experimental group, he monitored the display on which real-time analysis results were shown. He controlled the speed of lecture to allow students to catch up as much as possible. We evaluated the synchronization of the classroom; how many students opened the pages explained by the teacher. We counted the numbers minute-by-minute after setting a suitable delay.

Table 12.3 shows the ratio of synchronization of each group. For example, if we set the allowable delay to three minutes (i.e., if students opened the same page with the teacher within the three-minute delay), the synchronization ratio of the experimental group was 0.7661. The score was significantly different from the score of the control group. In other allowable delay settings, the synchronization ratios of the experimental groups were higher than those of the control group. We consider that such high synchronization was realized by the lecture speed control through the real-time feedback of classroom activities.

3.4 Learning Journal Analysis toward Teaching Analysis

3.4.1 Introduction

The practice of reflection is regarded as an important skill in higher education (Hatton & Smith, 1995). Reflection helps students see their learning activities objectively and deepen their understanding. Reflective writing is a representative activity of reflection. Through journal writing, students reflect on their process of learning and reconstruct the knowledge learned in the class. This allows students to see different aspects of learning activities and discover meaningful knowledge from it.

In reflective writing, it is common for students to record what they learn, what they wonder, and any other details regarding positive and negative perceptions of their learning experience. Therefore, from the viewpoint of learning analytics, researchers analyzed students’ journals for automatic evaluation of the engagement in reflective activities (Chen, Yu, Zhang, & Yu, 2016) and
Learning Support Systems Based on Cohesive Learning Analytics


While writing is very helpful for enhancing student learning, from the teacher's perspective, student journal writing is also useful for improving teaching. In a classroom, teaching and learning are very closely connected and thus, a student's negative experience could be caused by the teacher's technique or learning materials. Therefore, student journals are considered full of resources that enable teachers to objectively view their own teaching.

However, because of time constraints, it is challenging for teachers to grasp all useful information found in the journals. Reading journals is extremely time-consuming, and teachers may not be able to understand every journal, especially in the case of a large class. Therefore, computational assistance is crucial to a full utilization of student journals.

To this end, we consider two types of approaches: the method that supports teachers in reading journals and the automatic analysis method. The former gives us greater flexibility of analysis with a moderate amount of time required, and the latter provides easier understanding with less time for analysis. Neither way can fit every teacher's needs every time, so we need to appropriately combine both approaches.

In this article, we describe our proposed methods for both approaches based on weekly keywords. First, as the former approach, we describe a Web-based interactive reporting tool (Taniguchi, Okubo, Shimada, & Konomi, 2017), which makes it possible for teachers to explore student journal writings without losing important details. Second, we describe a method based on the latter approach that automatically extracts what students mention and how they think about them in their journals (Taniguchi, Suehiro, Shimada, & Ogata, 2017). As a case study last year, we applied our method on the journals for the course, “Information Science,” held for the first-grade students during the first semester 2016 at Kyushu University. We discuss the results in the rest of this section.

<table>
<thead>
<tr>
<th>Time</th>
<th>Control</th>
<th>Experimental</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 min.</td>
<td>0.4275</td>
<td>0.5174</td>
<td>0.0403*</td>
</tr>
<tr>
<td>3 min.</td>
<td>0.6598</td>
<td>0.7661</td>
<td>0.0033**</td>
</tr>
<tr>
<td>5 min.</td>
<td>0.7508</td>
<td>0.8599</td>
<td>0.0014**</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01
During our research, *weekly keywords* played an important role. Since our aim is to improve teaching, we want to understand what course-related topics were commonly discussed in the students’ journals writings. For the purpose, it is useful to measure how much each term used in journal writings is related to course contents. We assume that different topics are taught every week, and that students are told to write a reflection for each individual week. With this assumption, we can expect that the words related to the weekly topics will basically appear only in the journals for that particular week. Thus, using the week-specific nouns (*weekly keywords*) we can find the journal entries which are strongly related to the corresponding weekly topics.

Our method is based on the TF-IDF term-weighting measure (Salton & Buckley, 1988). The TF-IDF value of a word is defined as the product of term frequency (TF) and inverse document frequency (IDF). Since the IDF measures the document-specificity of words in a document set, we can measure the week-specificity by considering concatenated journal entries for a particular week as a single document. Table 12.4 shows extracted weekly keywords, where only the top ten words given the highest TF-IDF scores are shown. Comparing the keywords with actual topics, we can see that many keywords specific to course contents were successfully extracted.

We proposed a web-based interactive reporting tool for journal entries, which provides weekly keyword-based navigation of journal entries and summary plots of word usage patterns. The graphical user interface, shown in Figure 12.11, displays weekly keywords as entry points of journal exploration. The keywords are shown in ranked tables for each week according to their TF-IDF values, and allow to see how the usage of words varies from week to week.

In addition to that, there are also three interactive features. The first feature is a pop-up sentence view. When we hold the cursor over a word in the table, a popup appears and shows all the sentences containing the selected word in the corresponding week. This is helpful for us to understand in what context and how the word is used in real sentences.

The second feature is word highlighting, which emphasizes all occurrences of a highlighted word across the ranking tables. Since the ranking table shows the relative importance in a particular week, this feature makes it possible to see how the usage pattern of the word changes through a semester.

The last feature is grouping adjectives and verbs by their polarity, based on a dictionary. We can toggle the feature, and get positive and negative words in ranking tables colored with green and red colors, respectively. This is helpful in order to quickly assess journal contents from a sentimental point of view. Similar to this feature, we also show two types of plots in Figure 12.12. These
**Example results of weekly keywords extraction.** Weekly keywords are extracted from the journal entries for the course “information science.” From the 1st week to the 14th, actual topics taught in classes and extracted weekly keywords are listed. We can see topic-related words were extracted and general words were eliminated successfully.

<table>
<thead>
<tr>
<th>Week</th>
<th>Actual topics</th>
<th>Extracted weekly keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Positioning the lecture</td>
<td>Internet, orientation, university, search, binary, remaining, system, method, life, study</td>
</tr>
<tr>
<td>2</td>
<td>Introduction to information science</td>
<td>Morse code, signal, topic, quiz, homework, Japanese, thought, English, remaining, like</td>
</tr>
<tr>
<td>3</td>
<td>Information quantity and entropy</td>
<td>code, entropy, prefix, encoding, average, word, combine, length, symbol, multiple</td>
</tr>
<tr>
<td>4</td>
<td>Entropy</td>
<td>entropy, mutual, expectation, value, quantity, log, computation, with, information, condition</td>
</tr>
<tr>
<td>5</td>
<td>Channel coding</td>
<td>correction, error, automated, detection, Hamming, distance, communication, encoding, doodlebug’s pit, example</td>
</tr>
<tr>
<td>6</td>
<td>Cryptography, computer science</td>
<td>cipher, encryption, key, Caesar, mod, public, rsa, secret, decryption, high school</td>
</tr>
<tr>
<td>7</td>
<td>Computation, algorithm, time complexity</td>
<td>coin, Euclid, mutual division, balance scale, fake, algorithm, method, mathematics, gcd, rectangle</td>
</tr>
<tr>
<td>8</td>
<td>Midterm exam</td>
<td>exam, midterm, one question, perfect score, final, right answer, question, miss, two questions, effect</td>
</tr>
<tr>
<td>9</td>
<td>Stack and queue, bubble sort</td>
<td>notation, Polish, queue, stack, infix, sort, order, bubble, principle, first</td>
</tr>
<tr>
<td>10</td>
<td>Heap sort, merge sort</td>
<td>sort, heap sort, merge, tree, binary, comparison, algorithm, binary tree, drawback, practice</td>
</tr>
<tr>
<td>11</td>
<td>Bucket sort, binary search</td>
<td>sort, bucket, search, binary, search, binary, dictionary, Google, comparison, heap sort</td>
</tr>
<tr>
<td>12</td>
<td>Digital images</td>
<td>image, app, install, usage, download, exercise, Irfanview, rose, file, organism</td>
</tr>
<tr>
<td>13</td>
<td>Image processing, character recognition</td>
<td>image, recognition, processing, next week, letter, unistroke, edge, final, report, single stroke</td>
</tr>
<tr>
<td>14</td>
<td>Final exam</td>
<td>exam, final, small, perfect score, report, first semester, two questions, three questions, plan, final</td>
</tr>
</tbody>
</table>
charts make it possible for teachers to track the temporal changes in adjectival word usage, which in turn helps them identify topics interesting or difficult for students.

According to a short survey about our tool, responding teachers believed our tool does make it easier to read journals and identify any difficulties with their students.

When reading journal entries to acquire the students’ impressions, adjectives are the most descriptive words. Since adjectives modifying nouns is universal, we can find them by considering their co-occurrences (Schütze & Pedersen, 1994) with nouns in journal writings. Based on this idea, we proposed a method to find students’ impressions in pairs of weekly keywords and their associated adjectives.

We considered the co-occurrences within a sliding window and used Normalized Point-Wise Mutual Information (NPMI) (Bouma, 2009) to quantify the degree of association between two words. Briefly speaking, NPMI tells us about the dependency that occurs between two words just as a correlation coefficient does.

Furthermore, for easy understanding, we employed Nonnegative Matrix Factorization (NMF) as proposed in (Sra & Dhillon, 2006) to extract the collective impressions. NMF is a low-rank matrix approximation similar to a principal component analysis, and allows us to obtain a reduced matrix representation of relationships. We apply NMF to a modified NPMI matrix, whose elements are scaled so that the matrix contains only positive values. In order to automatically determine an optimal value for a parameter of ranks, we employed the stability analysis-based method proposed in Greene, O’Callaghan, and Cunningham (2014).

Table 12.5 shows an example list of weekly keywords most strongly associated with the adjective “difficult,” and the list of adjectives most strongly
associated with the weekly keyword, “cipher.” From the table, we can see what weekly topics are frequently mentioned in the student journals when students were writing about something difficult for them, and students had troublesome feelings regarding “ciphers.”

Table 12.6 shows an example result of abstraction of impressions by NMF. For example, the first group can be interpreted as the impression relevant to understanding concepts, and they are strongly associated with cryptography, algorithms, and information theory. In this example, we can interpret the result as the difficulty of these course topics were most important to many students compared to other topics. Such abstracted view of learning journals helps us to collectively understand the impressions of course contents.

In this chapter, we discussed the analysis of students’ reflective writings for teaching analysis. We described two examples of analysis adopting different approaches. We designed and implemented a web-based user interface for exploring student learning journals interactively, which facilitated reading the

<table>
<thead>
<tr>
<th>Target</th>
<th>Associated weekly keywords/adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>“difficult”</td>
<td>Irfanview (0.379), exercise (0.271), Japanese (0.227), infix (0.217), notation (0.199), Caesar (0.193), merge (0.117), comparison (0.117), encoding (0.117), code (0.103)</td>
</tr>
<tr>
<td>“cipher”</td>
<td>troublesome (0.247), detailed (0.208), fun (0.118), good (0.117), amazing (0.111), easy (0.088), interesting (0.068), difficult (0.022)</td>
</tr>
</tbody>
</table>
entries. The other proposal is a method that automatically extracts students’ impressions and abstract them for collective understanding. In future work, we can consider more extended analysis of journal writings, for example, for comparative analysis of different teachers. We can extract class-specific key-words and compare various teaching styles of teachers, which leads to learning journal-based teaching analysis.

### 3.5 Learner-Centric Notification Based on Learning Analytics

Next we discuss learner-centric notification mechanisms for learning analytics platforms based on our experiences with M2B, a campus-wide platform at Kyushu University. We argue for the use of commodity devices, such as smartphones, smart watches, personal computers, cameras, low-cost eye trackers, microphones and IoT appliances, so as to complement conventional online learning logs with the rich environmental, social, and physiological contextual data.

Assessment of learning in open-ended environments often focuses on learning processes and as such, quantifying learning outcomes may be difficult. Researchers have explored the potential of multimodal learning analytics to capture and analyze learning processes in project-based learning (Worsley, 2012) and other open-ended learning contexts (Blikstein, 2013). More recently, Worsley and Blikstein analyzed learning in a hands-on engineering design context, using speech, gesture, and electro-dermal activity data (Worsley & Blikstein, 2015). Oshima, Oshima, and Matsuzawa (2012) developed and evaluated the Knowledge Building Discourse Explorer, which captures and visualizes collaborative learning processes based on social network analysis.

<table>
<thead>
<tr>
<th>ID</th>
<th>Descriptive adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>difficult, easy, fun, troublesome, interesting, difficult, interesting, new, amazing, few</td>
</tr>
<tr>
<td>2</td>
<td>heavy, light, easy, interesting, few, involved, interesting, long, difficult, difficult</td>
</tr>
<tr>
<td>3</td>
<td>many, difficult, smart, big, well, difficult, interesting, rare, short, right</td>
</tr>
<tr>
<td>4</td>
<td>good, regretful, amazing, interesting, detailed, wonderful, happy, casual, terrible, near</td>
</tr>
</tbody>
</table>
techniques. However, their research narrowly focuses on dialogue rather than overall knowledge building processes.

Some learning analytics projects have explored proactive capabilities of data-driven environments based on prediction (Okubo et al., 2017b) and mobile alerts (Straumsheim, 2013), although they do not fully respond to the dynamic contexts of learners. Context-responsive notifications have been used to provide personalized feedback regarding energy consumption (Igarashi et al., 2016) and safety (Sasao et al., 2017). However, they do not explicitly support learning processes.

3.5.1 Designing Learner-Centric Notifications

In order to support learners effectively, we considered a co-design process of notification, in which teachers and learners design mobile and wearable notification tools together to enhance the everyday lives of learners with relevant learning opportunities. Moreover, as students use mobile and wearable devices to access pervasive learning opportunities, the system continuously generates various sensor data to enrich learning analytics. This allows for incremental optimization of notification mechanisms based on machine learning techniques.

We imagine that the co-design process of notifications can make learning adaptable to personal goals and visions. This can help go beyond the limits of adaptive learning systems that rely on data about what happened in the past. As multimodal notifications can be developed in co-design workshops involving different stakeholders (Sasao et al., 2017), we opt for an approach that encourages teachers and students to design notifications in face-to-face settings or by using asynchronous collaboration tools.

3.5.2 Smart Delivery of Notifications

In order for notification environments to be successful, systems should be able to trigger the right notifications at the right time in the right place in the right way, and assure the quality of contextual learning experiences in everyday life settings. We can extend and integrate existing mobile and situated crowdsourcing techniques (Goncalves et al., 2017), as well as machine learning approaches, to predict the right context that would trigger notifications to learners.

Delivering notifications in the right way requires the appropriate presentation of notifications for different learners in different contexts. This would involve using not only visual, auditory and textual alerts, but also ambient displays. Moreover, systems could notify teachers who have the skill to present
the notifications to their students effectively. As participatory sensing expands, the potential of pervasive sensing, participatory actuation, or participatory presentation of notifications in particular, can expand pervasive learning in everyday contexts.

4 Conclusion

In this chapter, we introduced the tools for supporting learning and teaching with the utilization of the learning logs that are stored in the integrated learning support system, M2B. The system was developed to enhance learning awareness and performance both inside and outside the classroom consistently, based on the fundamental investigation of the relationships between learning behavior data, self-regulated learning theory, and learning patterns outside the classroom that promotes learning student academic achievement.

For grasping an overview of learning activities through a course, we introduced the Active Learner Dashboard that enables teachers and students to understand the learning situation in the course. From the user questionnaire, we found that this tool is effective for changing the student’s learning awareness, however, a system that responds to each student individually is considered necessary. For this reason, we conducted an automatic summarization of learning materials to promote previewing and real-time analysis to understand students’ behaviors during class time, as well as a learning journal analysis to learn students’ situations of learning and self-reflection after a class. The tools and analysis that match the learning situation have positive effects on the improvement of the learning environment and students’ learning performance.

Finally, we discussed learner-centric notification mechanisms for learning analytics platforms. Providing timely feedback of knowledge obtained through learning analytics to students and teachers with appropriate designs and triggers is a further important task to be investigated in the future, as suggested by our research.

Acknowledgements

This work was partially supported by JST PRESTO Grant Number JPMJPR1505, “Research and Development on Fundamental and Utilization Technologies for Social Big Data” (178A03), the Commissioned Research of National Institute of Information and Communications Technology (NICT), and Kyushu University Interdisciplinary Programs in Education and Projects in Research Development, Japan.
References


