Learning state estimation method by browsing history and brain waves during programming language learning

Katsuyuki Umezawa[1], Tomohiko Saito[2], Takashi Ishida[3], Makoto Nakazawa[4], and Shigeichi Hirasawa[5]

[1] Shonan Institute of Technology, Kanagawa, 251-8511, Japan,
umezawa@info.shonan-it.ac.jp,
WWW home page: http://www.info.shonan-it.ac.jp/umezawa-lab/index.html
[2] Shonan Institute of Technology, Kanagawa, 251-8511, Japan
[3] Takasaki City University of Economics, Gunma, 370-0801, Japan
[4] Junior College of Aizu, Fukushima, 965-8570, Japan
[5] Waseda University, Tokyo, 169-8555, Japan

Abstract. Various factors affect misstep learning, such as the quality and difficulty of the learning contents and the learner's proficiency level. A system for acquiring the browsing history at the time of learning has been proposed. However, it may be insufficient to refer only to the learner's browsing time. For example, when the browsing time is short, it can not be determined whether the learning contents were too easy for the learner or whether learning was abandoned because the learning contents were too difficult. Therefore, in this paper, we propose a method of determining the learning state of learners by simultaneously analyzing learning history information and brain wave information, not using history information and brain wave information individually. And we will show that the learning state for each learner will be able to be successfully estimated by our proposed method.

Keywords: learning analytics, simple electroencephalograph, brain wave, e-learning, learning log

1 Introduction

Methods for effectively utilizing web teaching materials [1] and integrating digital textbooks and e-Learning systems [2] have been proposed. In a similar manner, the authors developed and prototyped electronic teaching materials and evaluated them in classroom trials [3]. In addition, we developed a system for referencing the editing history of the learning materials [4]. We have applied the system to English education [5] and to programming education [6] .

The relationship between intellectual work and brain waves has been studied under the assumption that the β wave is strongly related to a person's mental state[7]. Antonenko et al.[8] present a review of seminal literature on the use of continuous EEG to measure cognitive load and describes two case studies on learning from multimedia[9] and hypertext[10] that employed EEG methodology to collect and analyze cognitive load data. Other researchers have found that the ratio of the α and β waves are effective in determining a person's state of mind [11] [12]. Memory performance has also been studied by using brain waves [13][14]. Furthermore, we found that the low-β-wave/low-α-wave ratio, where "low" means low frequency, increases as a person works on a difficult task [15] .

As mentioned abobe, a system for acquiring the browsing history at the time of learning has been proposed. It is also proposed to know the state of the learner by using the information of the brain waves. However, it may be insufficient to refer only to the learner's browsing time. For example, when the browsing time is short, it can not be determined whether the learning contents were too easy for the learner or whether learning was abandoned because the learning contents were too difficult. Therefore, in this paper, we propose a method of determining the learning state of learners by simultaneously analyzing learning history information and brain wave information, not using history information and brain wave information individually. The final objective of this work is to improve the understanding of the learner by grasping situation of the learner and presenting the teaching material of the difficulty suitable for that situation.

In section 2, we describe previous works, and in section 3 we describe the problem when using learning history information and brain wave information individually. We propose a new system and method to know the learning state of learners in section 4. In section 5 we describe experimental methods and results, and we analyze the experimental results in section 6. Section 7 summarizes the key points and mentions future work.

2 Previous works

2.1 Web-based learning log collection system

When somebody reads a PDF file on the Web, the information recorded in the reading log is only "somebody downloaded the PDF file," not "somebody looked at page x." In addition, although the actions of a visitor are recorded in the access log of the website, we cannot determine exactly "which PDF was looked at and how many times" because the action is not recorded when the cache function of the browser is used.

In our previous work, we use the log data collection function of Web-based learning support system [4]. When a page is opened or closed, this system logs such information as the content ID, the page number, the opened/closed date and time, and how many seconds the page was open. Learner authentication is achieved by incorporating the Moodle system. The Moodle system is set to give user information for the support system. In this way, the user name authenticated in the Moodle system is added to the last URL automatically. The user name is thereby output as a log with learning time.

2.2 Applicability of brain waves to learning

It has been empirically shown by studies in psychology and cerebral science that the corrugation of brain waves can be used as an index of the person's state of mind. In previous studies, the α wave and the β wave (obtained from the result of a discrete Fourier transform) have been used to determine a person's state of mind. The relationship between intellectual work and brain waves has been studied under the assumption that the β wave is strongly related to a person's mental state. Giannitrapani, for example, investigated the relationship between intellectual work and brain waves by measuring the brain waves of a person taking an intelligence test [7]. The low-frequency component of the β wave was found to be dominant during a reading and comprehension test, a mathematics test, and a diagram alignment test.

In addition, there are many researches that measured the cognitive load at the time of learning using EEG. Changes in the power of brain waves measured with EEG was also the method of choice in a study that examined cognitive processes in multimedia learning [9]. Antonenko et al. investigated the effects of leads on cognitive load and learning in an experiment who read lead-augmented and hypertexts of comparable conceptual difficulty using EEG [10]. Other researchers have found that the power spectra of the α and β waves, the ratio of the α and β waves in relationship to all brain waves, and the simple ratio of the α and β waves are effective in determining a person's state of mind [11] . It was also shown that the activity of a person's brain can be determined by measuring the α and β waves and estimating the value of β/α [12]. That is, as the processing load increases with task difficulty, so does the value of β/α.

In an experiment using a simple typing test with varying degrees of difficulty, we confirmed that the β/α ratio increases with task difficulty. Furthermore, we found that the low-β-wave/low-α-wave ratio, where "low" means low frequency, increases as a person works on a difficult task [15]. In another experiment in which the change in brain waves was measured as the examinees became used to a new task (assembling a robot using three-dimensional motion capture) we showed that, although the examinees became accustomed to the task in various ways, the low-β/low-α ratio of the examinees who reported that the task was easy fell gradually [16].

3 Problems

The following two cases can be considered that the time to browse learning materials is short. One is when the learning is completed in a short time because the contents are too easy for the learner. The other case is that the contents were too difficult and the learner abandoned browsing the learning materials. Also, there are several reasons why browsing time is long. One is that the content is difficult and takes a lot of time. Another reason is that the learner is not concentrating on learning but just learning materials are displayed on the screen and the learner does not think. In this way, it is difficult to grasp accurately the state of the learner only with the log information of the browsing history.

Therefore, in this research, we propose a method to comprehend learners' learning state by simultaneously analyzing learning history information and electroencephalogram information, not using history information and biometric information individually.

4 Proposal

4.1 Proposed system

We show the proposed system in the left side of Fig.1. The proposed system acquires browsing history information and brain wave information from an existing browsing history system and electroencephalogram measurement system. And it analyzes the learning state of learners. First, the learning log collecting part obtains browsing history information from the learning log collection system and stores it in storage. At the same time, the brain wave collecting part acquires the brain wave information from the brain wave collection system and stores it in the storage. Using both the stored browsing history information and brain wave information, the analysis part analyzes what learning state each user was in, and stores the analysis result in the storage.

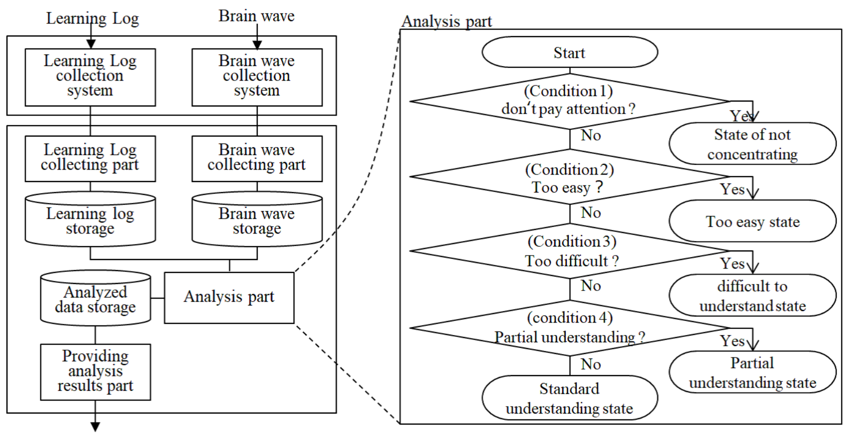

Fig. 1. Proposed system and algorithm

4.2 Proposed algorithm

A right side of Fig.1 shows the proposed algorithm for analysis performed in "Analysis part" of the left side of Fig.1. First, referring to the data representing the degree of attention by electroencephalogram measurement, it is estimated whether or not "you can not concentrate on learning (NC)" (condition 1). In other cases, it is estimated from the relationship between the brain wave contemplation degree and the teaching material browsing time whether "contents of learning is too easy (TE)" (condition 2). Further, it is estimated from brain wave contemplation degree and teaching material browsing time whether "contents of learning is too difficult (TD)" (condition 3). Finally, it is estimated whether "there is a part that can not be understood partially (PU)" (condition 4). If all conditions are not met, it is estimated to be "a standard understanding state (ST)".

Specifically, the determination is made as follows. First, in Condition 1, when the degree of concentration is low, it is estimated that this learner can not concentrate. Next, using Condition 2 and Condition 3, it is estimated whether it is too easy or too difficult. Since it is impossible to specify the degree of difficulty only by whether the learning time is long or short, it is possible to estimated whether or not the learner was thinking well by adding the brain wave information to the estimation expression. This makes it possible to distinguish between a case where the learner thinks carefully and takes long time (condition 3) and a case where the learner can easily understand and not take long time (condition 2). In addition, in the case where it is too difficult to abandon thinking, it is determined that the learner has not been able to concentrate already according to the condition 1. Lastly, in Condition 4, estimation of partial understanding is made according to the number of times of going back and forth on each page of the teaching material. Finally, if Condition 1 to Condition 4 do not apply, the learner is estimated to have a standard understanding.

5 Experiment

5.1 Experimental outline

"Matsudai Science Course" was held for students of Niigata Prefectural Matsudai High School and its neighboring high school students. And as part of this science course we conducted experiments. In this experiment, 18 high school students were taken as examinees. Assuming the e-learning, the students were allowed to browse slides (total of 8 pages) that explains the basics of C language that primary scholars should learn, and the brain waves at that time were measured.

The experiment was conducted in a quiet computer room.

This experiment has been approved by the Research Ethics Committee of Shonan Institute of Technology. In addition, we received signatures from students and parents of them concerning experiment participation.

5.2 Acquiring method of browsing history

In this experiment, we used the log collection system that we have already descrived in section 2.1. When we open or close a page, this system logs the information such as content ID, page number, opened/closed date and time, and how many seconds the page was open. Specifically, we set the Moodle system to give user information for this log collection system. In this way, the user name authenticated in the Moodle system is output as a log with learning time. The contents for the study are 12 slides about C programming language. After students finished study, we collected the logs for the time in second that students looked at each page.

5.3 Measurement method of brain waves

The EEG used for measuring the brain waves was a MindWave Mobile headset (NeuroSky, Inc.). The headset was connected to a ThinkGear Connector by Bluetooth, and the ThinkGear Connector communicated with a log-collection application by TCP/IP. The ThinkGear Connector is a driver provided by NeuroSky Inc. for communicating with the MindWave Mobile headset. As shown in [17], eight types of brain waves could be acquired, as shown in Table 1. A four-byte (unit-less) floating-point value was acquired for each type. In addition, MindWave Mobile headset can get eSense (i.e. Attention, Meditation) data that has been developed for NeuroSky devices specially. Attention and Meditation values were scaled between 1 to 100 and evaluation of the rating scale is as follows [18]: Measurements between 40 and 60 are considered as “neutral”. It is accepted as a little high between 60 to 80. It is accepted as high between 80 to 100. Values between 20 to 40 are considered as low. Values between 1 to 20 are shown to be very low. In this proposed algorithm, we use three types of data such as α_l, β_l, and Attention.

Table 1. Brain Waves That Could be Acquired

Type	Frequency (Hz)	Type	Frequency (Hz)
δ wave	0.5–2.75	low β (β_l) wave	13–16.75
θ wave	3.5–6.75	high β (β_h) wave	18–29.75
low α (α_l) wave	7.5–9.25	low γ wave	31–39.75
high α (α_h) wave	10–11.75	mid γ wave	41–49.75

5.4 Result of experiment

We collected a browsing history log of how many seconds the students viewed what page. Also, from the EEG, α wave, β wave, attention degree and meditation degree were measured at 1 second intervals (In the experimental simple EEG, in addition to the α wave and β wave, the degree of attention and meditation can be measured). Fig. 2 shows an example of part of log information that integrates the log of the students' browsing history and brain wave history. According to the reference [15], we considered the value of (low β wave)/(low α wave) to represent difficulty of task and decided to call it contemplation degree.

date	User ID	Page	Attention	Meditation	α_l	β_l
2016/08/18T11:30:21	ma001	3	56	41	1097	883
2016/08/18T11:30:22	ma001	3	70	23	138094	62256
2016/08/18T11:30:23	ma001	3	96	10	12529	22642
2016/08/18T11:30:24	ma001	4	100	1	3034	8763
2016/08/18T11:30:25	ma001	4	100	3	128468	7349
2016/08/18T11:30:26	ma001	4	97	13	12797	5164

Fig. 2. Example of learning log and brain wave log

6 Analysis

6.1 Result of analysis

Based on the logs obtained as a result of the experiment, the learning state of the learner was estimated by the proposed algorithm shown in the right side of

Fig. 1. For the conditions 1 to 4 of the proposed algorithm, we used the following calculation formulas for the present analysis:

$$\overline{AT_i} < 50 \qquad \text{(Conditioin 1)}$$
$$B_i + T_i < -40 \qquad \text{(Conditioin 2)}$$
$$B_i + T_i > 40 \qquad \text{(Conditioin 3)}$$
$$P_i > 6, \qquad \text{(Conditioin 4)}$$

where

$$\overline{AT_i} = \Sigma_j AT_{ij}/n_i \tag{1}$$
$$B_i = 100 \times (\overline{BW_i} - \overline{BW}) \tag{2}$$
$$T_i = TM_i - \overline{TM} \tag{3}$$
$$\overline{BW_i} = \Sigma_j BW_{ij}/n_i \tag{4}$$
$$\overline{BW} = \Sigma_i \overline{BW_i}/m \tag{5}$$
$$\overline{TM} = \Sigma_i \overline{TM_i}/m. \tag{6}$$

Here AT_{ij} is an attention of a student i at time j, BW_{ij} is β_l/α_l of a student i at time j, TM_i is reading time of a student i, n_i is a number of mesurement data of a student i, P_i is a number[6] of page return times of a student i, and m is the number of students.

Table 2 shows the estimated result of the learning state by the above formulas (Note that the questionnaire results on the right side of Table 2 are described in the next section.).

6.2 Discussion of results

In this section, we interpret the learning state of subjects estimated by the proposed algorithm against the questionnaire results carried out after the experiment. The questionnaire is as follows.

Q1: Do you think the learning materials were overall easy to understand?
Q2: Do you think the learning materials were difficult to understand in some places?
Q3: Do you think the learning materials were overall easy to see?

The answers to each question are as follows.

A1: Agree.
A2: Agree a little.
D2: Disagree a little.
D1: Disagree.

[6] If this value is large, it means that you are rereading it many times, it is considered partially unintelligible.

Table 2. Estimated results

ID	TM_i	T_i	$\overline{AT_i}$	$\overline{BW_i}$	B_i	$B_i + T_i$	P_i	Estimated result * (leftmost is valid)					Questionnaire		
													Q1	Q2	Q3
ma001	154	-39.50	66.60	1.70	34.69	-4.81	3					ST	A2	A2	A1
ma002	247	53.50	67.29	1.14	-21.50	32.00	11				PU		A2	D2	A1
ma003	225	31.50	43.19	1.45	9.90	41.40	4	NC		TD			D2	A2	A2
ma004	241	47.50	50.43	1.86	50.73	98.23	5			TD			A1	A2	A1
ma005	177	-16.50	57.21	0.94	-41.71	-58.21	4		TE				A2	A1	A1
ma006	231	37.50	44.06	1.37	1.95	39.45	7	NC			PU		A1	D1	A1
ma007	199	5.50	50.12	1.32	-3.55	1.95	0					ST	D2	A2	A2
ma008	134	-59.50	50.78	1.88	52.62	-6.88	0					ST	A2	A2	A1
ma009	98	-95.50	54.59	1.02	-33.49	-128.99	0		TE				A1	D1	A1
ma010	147	-46.50	66.97	1.16	-19.73	-66.23	0		TE				A1	D1	A1
ma011	168	-25.50	59.25	1.28	-7.27	-32.77	0					ST	A1	A1	A1
ma012	243	49.50	65.75	1.34	-1.53	47.97	0			TD			A1	A1	A1
ma013	219	25.50	79.24	1.36	0.86	26.36	8				PU		A1	D2	A1
ma014	211	17.50	59.81	1.47	11.93	29.43	0					ST	D2	A2	D2
ma015	283	89.50	30.25	1.26	-9.33	80.17	0	NC		TD			A2	A2	A1
ma016	158	-35.50	42.93	0.89	-46.34	-81.84	1	NC	TE				D2	A2	A1
ma017	133	-60.50	65.82	1.67	31.27	-29.23	7				PU		A1	D1	A1
ma018	215	21.50	46.52	1.26	-9.51	11.99	6	NC			PU		–	–	–
Average	193.50	0.00	55.60	1.35	0.00	0.00	3.1								

* NC: non-concentration, TE: too easy, TD: too difficult,
PU: partial understanding, ST: standard understanding.

The questionnaire result is shown in the right side of Table 2. The two students (ma009 and ma010) classified as TE are totally clear (Q1 is A1), and would not have thought that there was anything difficult to understand (Q2 is D1). In addition, those who classified as PU (ma002, ma006, ma013, ma017) answered "Disagree" or "Disagree a little" in response to the question that it was difficult to understand in some places (Q2 is D1 or D2). This can be interpreted as having difficulty to understand by going back to the page and understanding again because there were difficult parts. The interpretation of "partial understanding" may also be interpreted as "there are parts that can not be understood temporarily and understanding is deepened by page return". Students classified as TD (ma004, ma012) responded that there were places that were difficult to understand in some places (Q2 is A1 or A2) [7]. Also, if we don't consider the degree of attention, the student (ma016) will be classified as TE with a low degree of contemplation. However, this means that he can not pay attention on it and he can not contemplate it, so we can think that he was successfully classified by the proposed algorithm.

[7] Their answers to Q1 and Q2 are contradictory and can be said to be the limit of questionnaires that depend on subjectivity. In the questionnaire there was a contradiction, but with our algorithm they could be classified as TD. However, it seems that the material used is too simple from the questionnaire result. We think the experiment should be repeated with more difficult material.

7 Conclusion

The learning state for each learner was able to be estimated by this proposed algorithm. This enables us to grasp the situation of each student, so it will be possible to provide educational materials of difficulty suitable for individual students in real time in the future. In this time we made an estimation on the learning state throughout the learning materials, but we think that we can connect to the guidelines for preparing learning materials by making an estimation for each page of teaching materials. In addition, it is also necessary to establish a policy for setting thresholds and coefficients for Condition 1 to Condition 4. We also believe that it is necessary to conduct a questionnaire on the same items as the judgment result by the proposed method and obtain the correlation. Furthermore, integration with editing history in the case of learning for editing is also a future work.

Acknowledgment

In conducting this experiment, we received much cooperation from Mr. Masakazu Hasegawa of Matsudai High School of Niigata Prefecture. In addition, Mr. Kazuyuki Kido (representative of Waseda Matsudai Cooperation Association) gave us invaluable support such as coordination between the local community and the university and running of lectures. A part of this work was supported by JSPS KAKENHI Grant Number JP16K00491, JP17K01101, Special account 1010000175806 of NTT Comprehensive agreement Collaborative research of Waseda University Research Institute for Science and Engineering, and the grant as a research working group "ICT and Education" of JASMIN.

References

1. Ai Sugimura, Masahiro Ozaki, Saori Takeoka, and Yoshinori Adachi 2009. "The effective use of the Web teaching materials in class," (in Japanese) The Institute of Electronics, Information and Communication Engineers (IEICE) Technical report, ET, Vol. 108 (470), pp. 7–12.
2. Yasushi Suzuki November 2011. "The Effects of Seamless Combining Digital Textbooks and e-Learning," (in Japanese) Japan Universities Association for Computer Education, Journal of Educational Application of Information and Communication Technologies, Vol. 14, No. 1, pp. 31–35.
3. Katsuyuki Umezawa, Takashi Ishida, Manabu Kobayashi, and Shigeichi Hirasawa October 2013. "Effectiveness Evaluation of Practical Use of the Electronic Teaching Materials for University Education," (in Japanese) National Conference of JASMIN 2013 Autumn, Japan Society for Management Information, pp. 45–48.
4. Michitaka Aramoto, Manabu Kobayashi, Makoto Nakazawa, Michiko Nakano, Masayuki Goto, and Shigeichi Hirasawa March 2016. "Learning Analytics via Visualization System of Edit Record - System Configuration and Implementation," (in Japanese) 78th National Convention of Information Processing Society of Japan, Vol. 4, pp. 527–528.

5. Michiko Nakano, Michitaka Aramoto, Satoshi Yoshida, and Kei Koutou March 2016. “Learning Analytics via Visualization System of Edit Record - Application to English Writing Task: Error Gravity and Error Correction Time,” (in Japanese) 78th National Convention of Information Processing Society of Japan, Vol. 4, pp. 531–532.
6. Masayuki Goto, Kenta Mikawa, Gendo Kumoi, Manabu Kobayashi, Michitaka Aramoto, and Shigeichi Hirasawa March 2016. “Learning Analytics via Visualization System of Edit Record - Analytics Model Based on Edit Record and Evaluation Score Data for C-Programming Courses,” (in Japanese) 78th National Convention of Information Processing Society of Japan, Vol. 4, pp. 533–534.
7. D. Giannitrapani, 1988. “The role of 13-hz activity in mentation,” The EEG of Mental Activities, pp. 149–152.
8. Pavlo Antonenko, Fred Paas, Roland Grabner, and Tamara van Gog, “Using Electroencephalography to Measure Cognitive Load,” Educational Psychology Review, Vol. 22, Number 4, p.p. 425–438. 2010.
9. Gerlic, I., and Jausovec, N. “Multimedia: Differences in cognitive processes observed with EEG,” Educational Technology Research and Development, 47(3), p.p.5–14. 1999.
10. Antonenko, P., and Niederhauser, D., “The effects of leads on cognitive load and learning in a hypertext environment,” Computers in Human Behavior, 26, p.p. 140–150. 2010.
11. Hidetake Uwano, Kyoko Ishida, Yuko Matsuda, Shota Fukushima, Noboru Nakamichi, Masao Ohira, Ken-ichi Matsumoto, and Yasunori Okada 2008. “Evaluation of Software Usability Using Electroencephalogram - Comparison of Frequency Component between Different Software Versions,” (in Japanese) Journal of Human Interface Society, vol. 10(2), pp. 233–242.
12. Kouji Yoshida, Yuuta Sakamoto, Isao Miyaji, Kunihiro and Yamada 2012. “Analysis comparison of brain waves at the learning status by simple electroencephalography,” KES’2012, Proceedings, Knowledge-Based Intelligent Information and Engineering Systems, pp. 1817–1826.
13. Fumiyasu Hirai, Kouji Yoshida, and Isao Miyaji 2013. “Comparison analysis of the thought and the memory at the learning time by the simple electroencephalograph,” (in Japanese) Multimedia, Distributed, Cooperative, and Mobile Symposium (DICOMO 2013), pp. 1441–1446.
14. Fumiyasu Hirai, Kouji Yoshida, and Isao Miyaji 2014. “Trial of the EEG state feed-back learning system at the time of the memory work by the simple electroencephalograph,” (in Japanese) Multimedia, Distributed, Cooperative, and Mobile Symposium (DICOMO 2014), pp. 633–638.
15. Katsuyuki Umezawa, Takashi Ishida, Tomohiko Saito, Makoto Nakazawa, Shigeichi Hirasawa December 2016, “A judgment method of difficulty of task for a learner using simple electroencephalograph,” (in Japanese) Information Processing Society of Japan (IPSJ) SIG Technical Report.
16. Katsuyuki Umezawa, Takashi Ishida, Tomohiko Saito, Makoto Nakazawa, Shigeichi Hirasawa February 2017, “A study on transition of the brain waves when executing a task using simple electroencephalograph,” (in Japanese) Information Processing Society of Japan (IPSJ) SIG Technical Report.
17. ThinkGear Serial Stream Guide
http://developer.neurosky.com/docs/doku.php?
id=thinkgear_communications_protocol (2017/3/30)
18. MindWave Mobile: User Guide August 5, 2015.