Extraction of Characteristic Answering Behavior Using Handwritten Log Data

Junya ATAKE[a]*, Taito KANO[a], Kohei NAKAMURA[a], Chia-Yu HSU[a], Izumi HORIKOSHI[b] & Hiroaki OGATA[b]
[a] *Graduate School of Informatics, Kyoto University, Japan*
[b] *Academic Center for Computing and Media Studies, Kyoto University, Japan*
*atake.junya.86t@st.kyoto-u.ac.jp

Abstract: We extracted learners' characteristic answering behaviors from handwritten process log data and investigated whether learners' situations could be inferred based on these characteristics. The result showed we were able to extract several characteristic answering behaviors, such as stopped pen stroke and late start. Furthermore, we examined the learners' situation for each feature in the actual answering process. The results revealed that several characteristic answering behaviors indicated situations such as learners' stumbling or giving up. These results imply that handwritten process log data can allow teachers to capture learners' situations and support teachers' interventions.

Keywords: Handwriting, log data, answering behavior, learning analytics

1. Introduction

In Japan, under the Global and Innovation Gateway for All (GIGA) school project promoted by the Ministry of Education, Culture, Sports, Science, and Technology, a digital device was provided to each learner, and educational log data were accumulated. A typical digital device used in the classroom is a handwriting tool that allows learners to write directly on the devices. One advantage of using such tools is that learning logs, which represent the writing process, answer time, and handwriting, can be collected and analyzed. It is important for teachers to analyze the learners' answers to provide appropriate instructions using the handwritten process log data. However, it is difficult for teachers to analyze the learners' answers and understand the situation within a limited time.

Several studies have been conducted on learners' handwriting processes, such as the detection of answer stuck points from character recognition and the pen-stroke, the time interval of one stroke in handwriting (Iiyama et al., 2017), and the detection of deleted or revised strokes using support vector machines (Maeda et al., 2015). These methods make it easier for teachers to understand learners' stuck points, deletions, and corrections. However, both studies focused on their answering contents and did not focus on their answering behaviors or situations.

This study extracted the learners' characteristic answering behaviors from the handwritten process log data, and investigated whether inferring the learners' situation is possible. "Characteristic answering behaviors" refer to "Interval of stopping Pen-stroke," and "Eraser usage," and so forth, which might reflect that learners spend time reflecting on and correct their mistakes. In addition, we attempted to identify the learners' situations based on the extracted characteristic behaviors. We set the following research question for this study.

RQ: Can we infer the learners' situations by extracting their characteristic answering behaviors from the handwritten process log data?

We expect that the inference of learners' situations from the features extracted in this study can contribute to automatically detecting learners' situations from the data, helping teachers understand their situations, and providing them with appropriate support.

2. Methods

2.1 Learning and Evidence Analytics Framework (LEAF)

We used handwritten process log data collected from the LEAF system for analyses. The LEAF system is a learning analytics platform that supports teaching and learning by analyzing accumulated educational data (Ogata et al., 2018). It consists of three main sub-tools: BookRoll, Moodle, and LogPalette, and log data are stored in the Learning Record Store (LRS). BookRoll is an e-book reader through which teachers upload learning materials to their learners. Moodle is a Learning Management System (LMS) used to manage courses and materials. LogPalette is a learning analytics dashboard that visualizes the interactions of learners. Learners can use the handwriting tools to answer the learning materials that the teacher has registered on BookRoll, and the teacher can replay the learners' answering processes on LogPalette.

2.2 Data and Analyses

This study's handwritten process log data were from the context in which a teacher asked learners to answer questions in a Japanese junior high school mathematics class. The learners started answering the questions simultaneously when instructed by the teacher, and answered three questions. Handwritten process log data includes the time (Timestamp) when each stroke was added and its position (x-coordinate, y-coordinate). We conducted two analyses of the data, as shown in Figure 1.

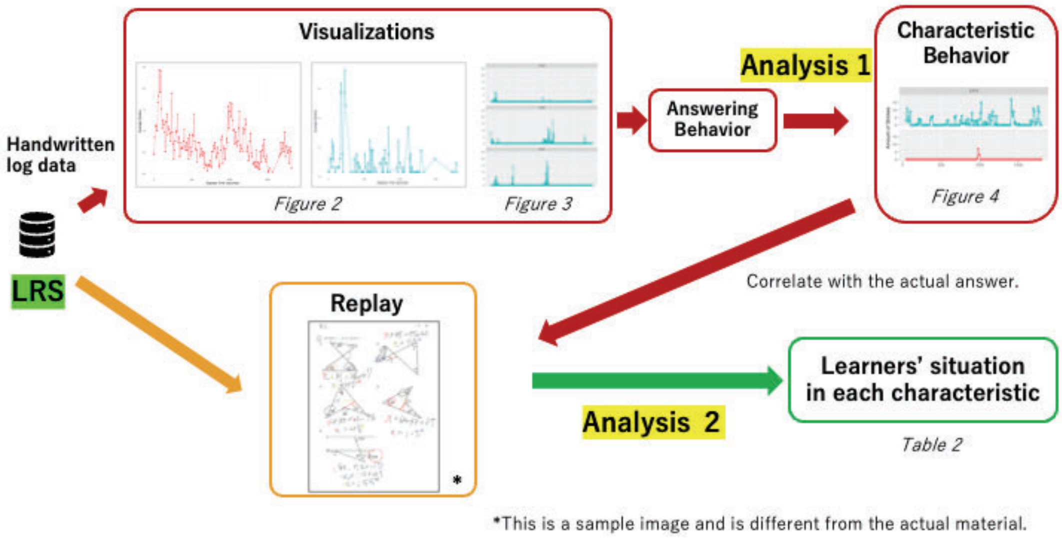


Figure 1. Analysis Flow.

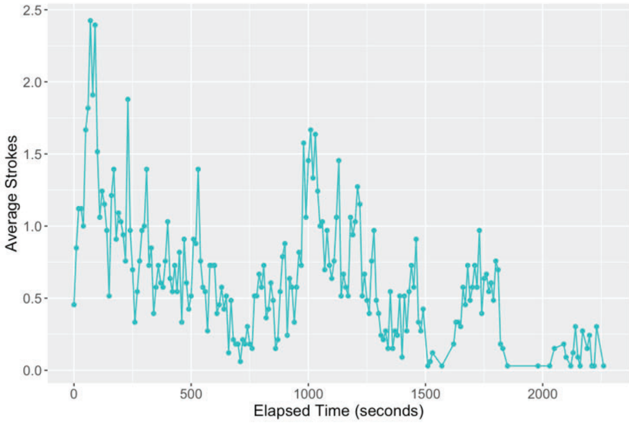
In Analysis 1, we visualized the log data as time-series graphs and extracted the characteristic answering behaviors. In these graphs, the x-axis shows the elapsed time, and the y-axis shows the number of strokes and eraser usage per 10 seconds.

In Analysis 2, we examined what kinds of situations could be observed in each characteristic answering behavior extracted from Analysis 1. We observed the actual answering process using the replay, and then examined whether there was a relationship between each characteristic and the learners' situations.

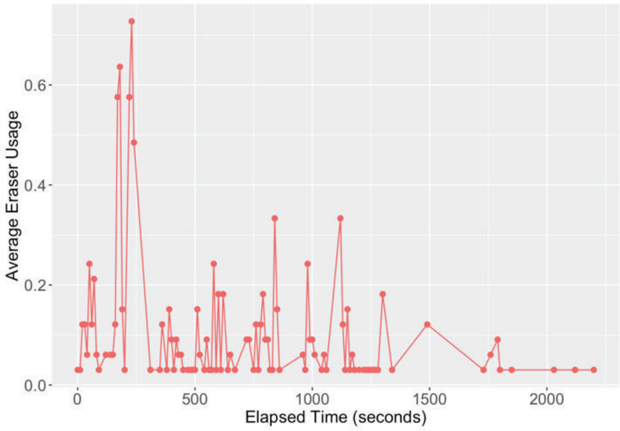
3. Results and Discussion

3.1 Analysis 1: Visualizing Log Data and Extracting Characteristic Answering Behaviors

First, we plotted the average number of strokes (Figure 2. (a)) and erasers used by all learners (Figure 2. (b)). We were able to ascertain the trends in the average number of strokes and the amount of eraser usage for each learner. Three peaks were observed in Figure 2. (a) for the average stroke. These peaks imply the culmination of the answers to the three questions.



(a) The Average Number of Strokes for All Learners.



(b) The Average Number of Using Erasers for All Learners.

Figure 2. Average Number of Strokes and Erasers.

Next, we plotted the answering behavior of all the learners (Figure 3) and extracted the characteristic answering behavior. Six characteristic features were identified (Figure 4).

Figure 4. (a), (b), (c), and (d) shows the behaviors characterized by how much time the strokes were stopped when the learners answered the questions. Specifically, Behavior A indicates the learners stopped the strokes for more than 50% of the answering time, Behavior B indicates the learners stopped for more than 80%, and Behavior C indicates the learners stopped for less than 20%. In addition, Behavior D indicates the learners paused and continued several times when answering. Every time they paused, the strokes were stopped for more than five minutes.

On the other hand, Figure 4. (e) shows the behavior characterized by the amount of using erasers. Behavior E indicates the learners used erasers with the amount above the average. Finally, Figure 4. (f) shows the behavior (Behavior F) that the learners waited for more than five minutes and started to answer after the teacher gave the instruction. It is noted that these features were not mutually exclusive and might overlap.

In this manner, we were able to extract some characteristic answering behaviors from the results. These features may reflect the learners spent time thinking and corrected mistakes. However, it was difficult to identify the learners' situation from these results alone, and it proved necessary to observe the state of the actual answers. Therefore, a more detailed analysis was performed in Analysis 2.

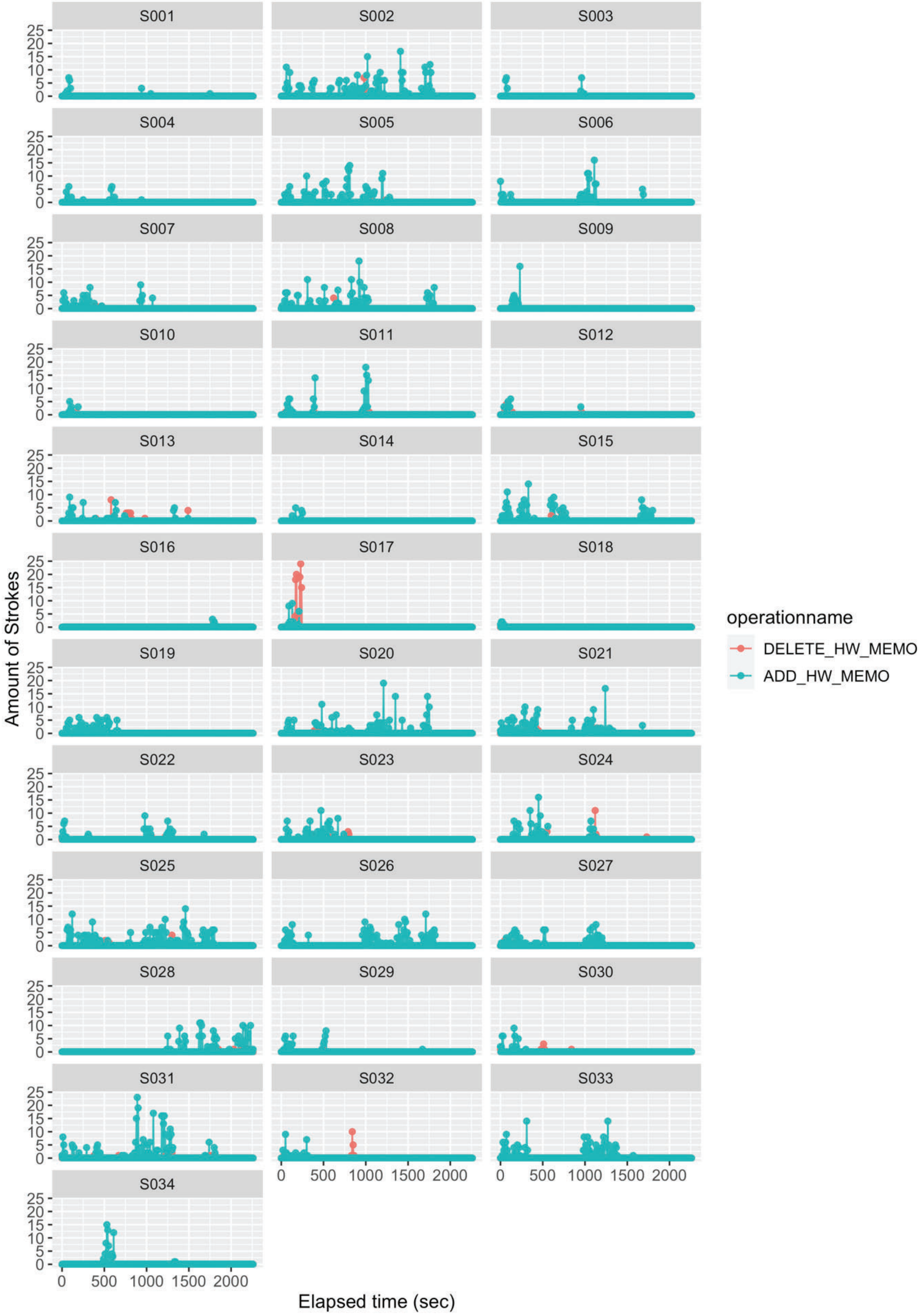
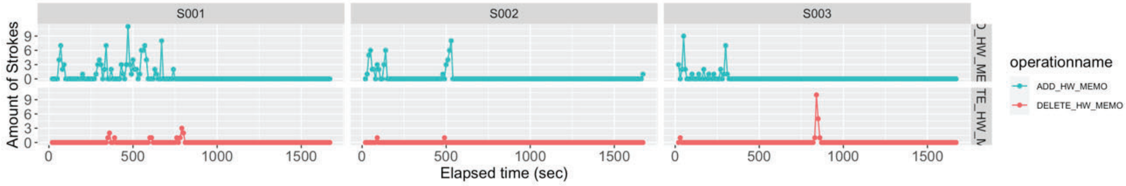
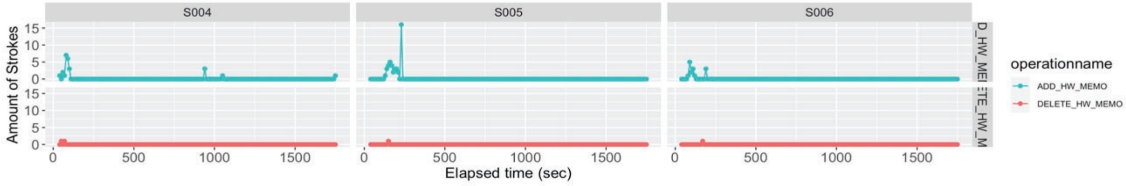


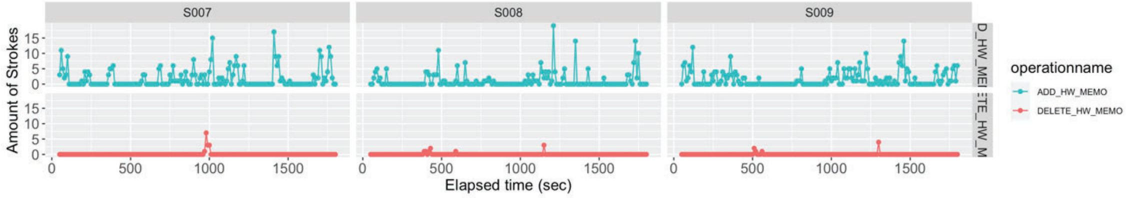
Figure 3. Answering Behaviors for Each Learner.



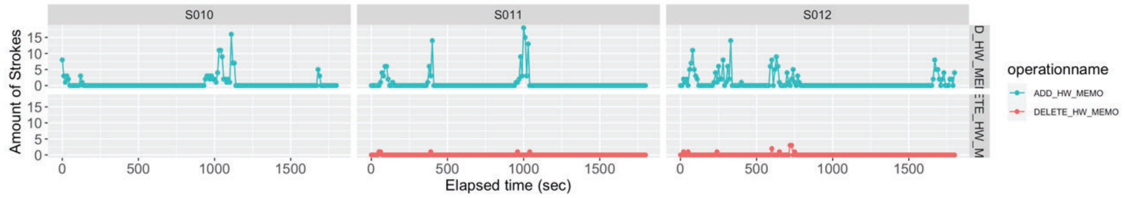
(a) *Behavior A: Stopped Pen Stroke Over the Half.*



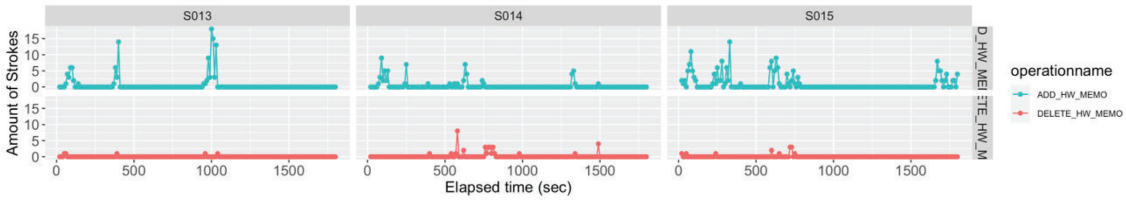
(b) *Behavior B: Mostly Stopped Pen Stroke.*



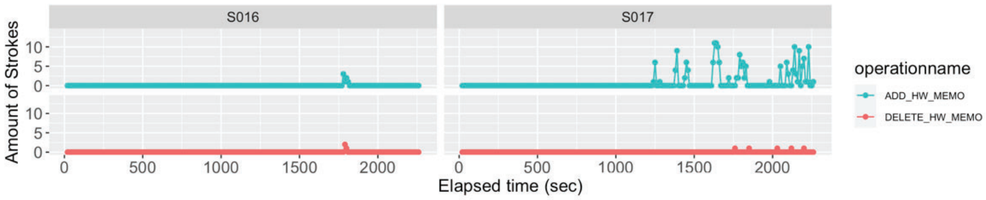
(c) *Behavior C: Few Stopped Pen Strokes.*



(d) *Behavior D: Two or More Points Where Pen-stroke Stopped for More than 5 Minutes.*



(e) *Behavior E: A Large Amount of Using Erasers.*



(f) *Behavior F: Late Start.*

Figure 4. Characteristic Answering Behaviors.

3.2 Analysis 2: Grasping the Situation of the Learner in Each Characteristics from the State of the Actual Answers

In Analysis 2, we replayed the actual handwriting process and observed what kinds of learners' situations could be understood from the features extracted from Analysis 1. We were able to identify the learners' situations, as shown in Table 1.

Table 1. *Characteristic Answering Behaviors and Learners' Situations*

Behavior	Feature (The number of learners)	Learner situation (The number of learners who matched to the situation)
A	Stopped pen stroke over the half (n = 5)	Give up on an answer halfway (n = 5)
B	Mostly stopped pen stroke (n = 6)	Give up on an answer halfway (n = 6)
C	Few stopped pen strokes (n = 6)	(1) Keep hands moving (n = 6) (2) Tend to get correct (n = 6)
D	Few stopped pen strokes (n = 5)	Answer only certain questions (n = 5) Give up on an answer halfway (n = 5)
E	A Large amount of using erasers (n = 5)	Give up on an answer halfway (n = 5)
F	Late start (n = 2)	No ideas (n = 1) Have trouble opening materials (n = 1)

The results indicate that some features can suggest situations such as learners' stumbling or giving up. These features have the potential to help the teachers understand learners' situations and provide support interventions.

4. Conclusion & Future Work

In this study, we set the research question as "Can we infer the learners' situation by extracting the learners' characteristic answering behaviors from the handwritten process log data?" To answer this question, we visualized time-series data to extract characteristic answering behaviors, and analyzed the answering process for each of them. From the results, we identified some features such as stopped pen strokes and eraser use. Additionally, some features suggest unique situations for learners. Although further research is required, we expect that these results are not confined to mathematics courses. In future research, it would be conceivable to support teacher interventions using these features. If the learner's situation is automatically identified from the data, it is possible to assist the teacher in understanding the learner's situation during the class.

Acknowledgements

This study was partially supported by NEDO JPNP20006, JSPS KAKENHI JP22K20246, and JP23H00505.

References

Iiyama , M., Nakatsuka, C., Morimura , Y., Hashimoto, A., Murakami, M., & Mino, M. (2017). Detecting answer stuck point using time intervals of pen strokes. *Transactions of Japanese Society for Information and Systems in Education, 34* (2), 166-171.

Maeda, Y., Miura, M., & Kobayashi, Y. (2015, August). Detecting erase strokes for understanding learning states from online handwritten notes. In *Proceedings of Informatic Education Symposium* (pp. 109-116).

Ogata, H., Majumdar, R., Akçapınar , G., Hasnine , M. N., & Flanagan, B. (2018). Beyond learning analytics: Framework for technology-enhanced evidence-based education and learning. In *International Conference on Computers in Education (ICCE 2018)* (pp. 486-489).