Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Gaze- and Semantics-aware Learning Material to Capture Learners' Comprehension Processes

Akio OKUTSU[a*], Yuki HAYASHI[b] & Kazuhisa SETA[b]
[a]*College of Sustainable System Sciences, Osaka Prefecture University, Japan*
[b]*Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Japan*
*okutsu@ksm.kis.osakafu-u.ac.jp

Abstract: Learners' comprehension processes and their states of knowledge can be used to analyze and facilitate learning. The purpose of this study is to develop a system that records the learning processes to estimate the learners' comprehension processes with respect to the target domain-knowledge and state of knowledge from their gaze on the screen when viewing the learning material. Gaze is often interpreted as a part of one's thinking processes. In this study, we developed learning material that records learners' learning processes by measuring their gaze to determine the semantics they tried to understand. In addition, we created an authoring system for developing the learning material.

Keywords: Gaze behaviors, semantic network, learning material, comprehension processes

1. Introduction

It is essential for learners to recognize their own state of knowledge and to self-regulate their processes of understanding in order to learn as effectively as possible (Mayer, 2014). For example, when learning from a textbook, learners need to establish semantic connections with prior knowledge (Chi, & Wylie, 2014). However, these kinds of metacognitive techniques can be difficult for immature learners to acquire (Kayashima, Inaba, & Mizoguchi, 2008). To encourage the use of these techniques, a learning support system must identify the learners' comprehension processes and state of knowledge in detail. Then the system can provide the learners with adaptive feedback and may also reveal the relationship between learning style and academic achievement. Using this analysis, the system may be able to estimate the comprehension level and provide information such as posing questions to encourage understanding and drawing the learners' attention to information they need to pay further attention to. The significance of such learning support is that it is suitable to each individual's learning process. In contrast, conventional learning support systems encourage learners to engage in different problem-solving situations to diagnose their state of understanding.

Many studies have focused on gaze as a way to understand the implicit nature of a learners' thinking processes. Gaze measurement has been used effectively in various cognitive processing analyses (Ohno, 2002) because gaze behaviors reflect a part of the thought processes without interfering with higher-order human cognitive processing. Furthermore, gaze measurement can be combined with other analytical methods and has been recognized as a promising method to approach cognitive and metacognitive learning processes that cannot be captured by learners' self-reported cognitive judgments (Roderer & Roebers, 2014; Hayashi, Seta, & Ikeda, 2018).

Several studies have used gaze behaviors in learning contexts to reveal the characteristics of learners. Antonietti, Colombo, and Nuzzo (2015) found that the frequency of a learner's attention indicates what they consider important in the learning material, while the length of their attention indicates their interest in the information, and exceedingly long attention indicates difficulties in cognitive processing. Mason, Pluchino, and Tornatora (2013) reported that the sequence in which information is presented in text and images reveal the cognitive processes of integrating with prior knowledge. Furthermore, the order in which learners pay attention to a text and the corresponding image may indicate the learners' cognitive processes of constructing organized knowledge and mental models.

A high frequency of close attention by gazing back and forth between the text and image may be related to academic achievement (Mason, Pluchino, & Tornatora, 2015).

Many studies have attempted to capture learners' thought processes or higher-order cognitive processing using gaze information on learning materials; however, these studies only analyzed characteristics of gaze behavior and comprehensive states and processes from visual components (e.g., text, figures, tables) of the information. The studies lack a framework to clarify the characteristics based on the specific knowledge content (semantic structure) embedded in the learning materials. The researchers gave their semantic interpretation of learners' gaze behaviors on the learning material manually, but this is not practical. Therefore, in order to accumulate further academic knowledge on gaze analysis, it is necessary to build a framework that can handle the semantics of knowledge content in addition to the visual components.

We previously developed gaze- and semantics-aware learning material to adaptively pose questions on the basis of the learner's attention. The material can be used to examine the correlation between the learner's gaze on the learning material and the semantics of the information (Muroya, Seta, & Hayashi, 2021). To make this method applicable to a variety of learning domains, in this paper, we propose an authoring system for creating gaze- and semantics-aware learning materials. Then, using the developed learning materials, we measure learners' gaze information on science-related learning material and investigate the system's capability to analyze and estimate the learners' learning processes using the semantics of the knowledge content. For this purpose, we use learning material containing contradictory information, which would cause cognitive conflicts during the learning processes if a learner tries to integrate knowledge rather than merely store it. The characteristics of measured gaze information on the contradictory knowledge may be differentiated by the undertaken cognitive processes. Thus, we investigate the system's capability of measuring such characteristic gaze information concerning knowledge content and capturing the degree of schema formulation.

2. Functional Requirements

To support learners constructing well-organized knowledge in textbook-based learning, this study focuses on the learners' gaze on the learning material to estimate their comprehension processes and knowledge state. Our final goal is to build a framework for interpreting gaze behaviors and estimating the learner's knowledge state by representing the semantic relations of the learning materials in the system.

A semantic network is a well-known knowledge representation method that defines semantic relations between concepts using nodes (concepts) and links (relations). The network enables semantic connections between concepts in the learning material to be represented in a machine-readable manner, including information that is not explicitly mentioned on the textbook but should be learned in order to construct well-organized knowledge.

We establish a set of knowledge in the semantic network for a particular area of the learning material as the area of interests (AOIs) in advance. This results in gaze- and semantics- aware learning material as the system can detect the semantics of the information the learner is gazing at based on the eye-tracking device. When used with such learning materials, the system can record learning processes, such as who pays how much attention to which information or the order in which the learner looks at the information, and stores this data as a learning log. By analyzing the learning log, it may be possible to trace the comprehension process and estimate the state of knowledge of the learner.

The following three functional requirements must be satisfied for this system to be feasible:

- **Requirement 1.** *An authoring support function for creating gaze- and semantics-aware learning materials*: In order for the system to trace the comprehension processes from the learners' gaze behaviors, the system must have the information of what knowledge is represented at what area on the learning material. Therefore, an authoring function is needed which enables the author to create semantic networks for the target learning material and set AOIs on the material corresponding to the semantic structure of knowledge, thereby creating gaze- and semantics- aware learning materials.
- **Requirement 2.** *Functions for measuring and recording gaze information during learning*: Semantics-aware gaze measurement and recording functions need to be used in conjunction with the

learning materials stated above. The functions would capture gaze information such as the learner's gaze time and the order in which the learner focused on the content, in addition to the media properties they focused on (text, figures, tables, etc.) and record them as a learning log.

- **Requirement 3.** *Learning support functions based on estimated learner's knowledge state by learning records*: Support functions are needed based on estimating the learner's comprehension process and the state of knowledge from the learning log. Then the system needs to provide adaptive feedback, such as presenting information that may be not fully understood or guiding the eyes on the learning material to promote structural understanding (e.g., highlighting the learning contents).

In this paper, we aim to fulfill the Requirements 1 and 2, which are needed to realize Requirement 3 in the future.

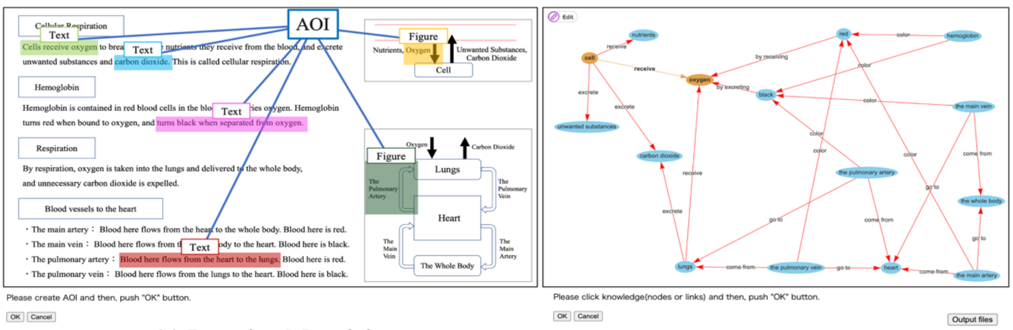
3. Proposed Systems

In this study, we developed two systems, an authoring system for gaze- and semantics-aware learning materials and a semantics-aware gaze measurement system to satisfy Requirements 1 and 2 described in the previous section.

3.1 Authoring System for Gaze- and Semantics-Aware Learning Materials

Figure 1 shows the interface of the developed authoring system. The system is implemented as a web application using JavaScript and HTML and runs in widely used browsers such as Chrome, Safari, and Edge. The interface consists of a learning material area (Fig. 1(a)) and a semantic network editing area (Fig. 1(b)).

The author uses the system by inputting the image file of the learning material (e.g., png format) and the file of the corresponding semantic network (xml format) that represents the semantic structure of knowledge. In the semantic network area, the author can freely edit and extend the displayed semantic network (e.g., create/delete nodes and links) accordingly for the learning material. In the learning material area, the author can drag the mouse to create a rectangular area indicating the AOI that captures the target of the learner's gaze from the eye-tracking device. Then the author selects the corresponding semantic structure of the knowledge to the AOI by clicking nodes and links in the semantic network editing area. After the selection, the system requires the author to select the media property of the AOI from 'Sentence,' 'Figure,' or 'Table.' By repeating these operations, a semantic network is linked to several areas in the learning material. Here, the system allows the author to associate different media in the learning material with the same knowledge in the semantic network. For example, if the author creates AOI on the text about "Cells receive oxygen", the corresponding semantic structure of the knowledge ("cell" node, "oxygen" node, and "receive" link) should be chosen in the semantic network, and 'Sentence' should be selected as the media property. On the other hand, if the author creates AOI on the area in a figure about the same information, the same semantic structure of the knowledge ("cell" node, "oxygen" node, and "receive" link) should be chosen in the semantic network, while 'Figure'



(a) Learning Material (b) Semantic Network

Figure 1. Authoring System for Gaze- and Semantics-aware Learning Materials.

should be selected as the media property.

Finally, the system outputs two files, the final state of edited semantic networks and AOI information that includes the ID of each AOI, position in the learning material, size, linked knowledge, and media property.

3.2 *Semantics-Aware Gaze Measurement System*

Figure 2 shows the developed gaze measurement system. The system is implemented as a C# form application and works with a screen-based eye tracker (Tobii Eye Tracker 4C with analytical license) to measure the learner's gaze behaviors. The system manages the learning material layer displayed to the learner (Fig. 2(a)) and the semantic network layer (Fig. 2(b)) that connects each AOI to the corresponding semantic structure of the knowledge.

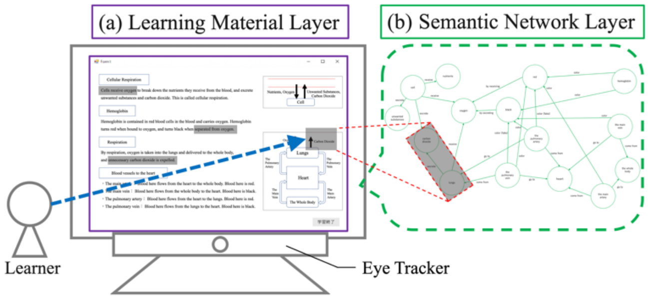


Figure 2. Semantics-aware Gaze Measurement System.

Before learning, the system requires the following three files to be input: the image file of the learning material (e.g., png format), the semantic network file (xml format), and the AOI information file (xml format) output by the authoring system described in Section 3.1. Then, the image of the input learning material appears on the interface. Although AOIs are set on the learning material on the basis of the input file (i.e., shaded area in Fig. 2(a)), the AOIs are invisible on the interface so as not to interfere with learning.

Learners carry out their learning using the learning material. During the learning processes, the system monitors the learner's gaze behavior, whether the gaze falls within certain AOI regions in each frame, and detects the corresponding knowledge in the semantic network. The information measured by the system is saved as a learning log in which each line includes the ID of the gaze target AOI, gaze state (start/stop), timestamp (ms), corresponding knowledge, and media property.

4. Experimental Study

To evaluate the feasibility of the proposed framework, we conducted the experimental study whether the developed systems fulfill the functional requirements 1 and 2 and confirm its potential. We created a learning material about blood circulation taught in junior high school science classes. The learning material included text and figures described in Japanese (Fig. 3(A)). We prepared 15 concepts and 34

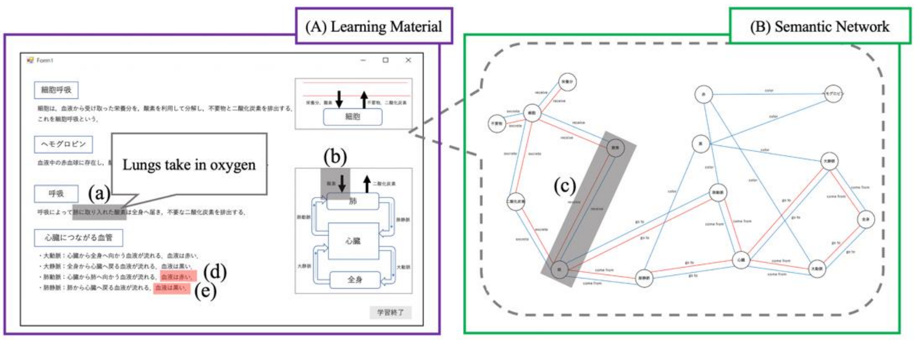


Figure 3. Learning Material and Its Semantic Network used for Learning.

links as the semantic network and set 29 AOIs using the authoring system. In the semantic network (Fig.3(B)), the blue and red links indicate the media property of the corresponding set of AOIs (blue: text, red: figure). For example, if the learner is gazing at the text "lungs take in oxygen (Fig. 3(A-a))" in the learning material, the system can detect that the learner is paying attention to the corresponding knowledge (Fig. 3(B-c)). The system distinguishes each AOI from the media property even if the related knowledge is the same. Thus, when the learner is gazing at the area in a figure about the information "lungs take in oxygen" (Fig. 3(A-b)), the system can grasp the corresponding knowledge (Fig. 3(B-c)) and record it separately from the text (Fig. 3(A-a)).

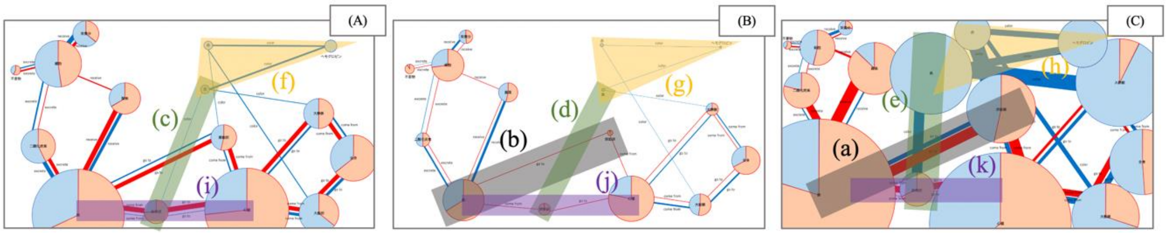


Figure 4. Visualization of Learning by each Learner.

In this experiment, the learning material includes incorrect information about the color of blood flowing through pulmonary arteries and veins as the text contents in the learning material (Fig. 3(A-d) and 3(A-e)). The objective of including incorrect information is to investigate whether the difference appears in gazing behaviors when the learner notices and causes cognitive conflicts from the inconsistent information.

Three junior high school students (learners A, B, and C) participated in learning using the semantics-aware gaze measurement system. We did not tell the learners that the material included incorrect information as mentioned previously.

Figures 4(A), 4(B), and 4(C) represent the visualized semantic network based on learning logs that recorded the learning processes of the three learners, respectively. The size of the circle (concept) and the thickness of the line (link) represent the total amount of time spent paying attention to the corresponding AOIs on the learning material. Here, the blue nodes and links indicate the rate of knowledge focused on textual information and the red ones represent the same for graphic information (i.e., figures on the learning material).

The results indicate that learner C spent the longest time among the three learners and paid attention to a wider variety of information. In addition, we observed the different tendencies of learners' focused media property. For example, while learner C paid attention to the information "the pulmonary artery is connected to the lung" in both the textual and graphic information (Fig. 4(C-a)), learner B only paid attention to the graphic information and the attention time was quite shorter than that of learner C (Fig. 4(B-b)).

In the post-learning interview, only learner C commented that she noticed the inconsistent knowledge (shown in Fig. 3(A-d) and 3(A-e)). The nodes and links corresponding to inconsistent information on the semantic network are "the color of the blood flowing through the pulmonary veins (Fig. 4(A-c), 4(B-d), 4(C-e))," "the color change of hemoglobin (Fig. 4(A-f), 4(B-g), 4(C-h))," and "the way the pulmonary veins connecting to the lungs and heart (Fig. 4(A-i), 4(B-j), 4(C-k))." From the three visualization diagrams, we found that learner C paid more attention to the inconsistent information than the other two learners. In addition, we found three characteristic gaze events from the timestamp information in learner C's log: (i) the ratio of total gazing time to the inconsistent information was larger than that of the other two learners, (ii) after she paid attention to the incorrect information, she gazed at the related information and (iii) she paid attention to both textual and graphic information represented the same knowledge. The ratio of gazing time (i) suggests that the learner may have been facing a cognitive conflict when she noticed the inconsistent information. She tried to structurally understand the learning contents by comparing and integrating them with her knowledge structure. To resolve the cognitive conflict, the learner would attempt to check the semantic connection between the inconsistent information one at a time, as suggested by (ii). (iii) suggests that the learner could recognize that the text and figures contained the same information, and she processed information systematically by contrasting the text and figures.

5. Concluding Remarks

In this study, we confirmed the proposed framework captures characteristics of each learner at the semantic level to some extent. It has potential towards learning contents oriented active learning analytics and supports. That is, in previous studies, the researchers manually analyzed the semantics of the knowledge in the learning material against information derived from gaze, such as cognitive and metacognitive processes of learners (e.g., Mudrick, Azevedo, & Taub, 2019) and the learning styles. In contrast to the conventional analysis, we proposed an authoring system for creating the gaze- and semantics-aware learning materials (Requirement 1) and a measurement system (Requirement 2) that can automatically capture several semantic relationships of information represented in the learning material (e.g., contents containing the same information, other semantically connected information, and inconsistent information).

Learning is a complex and implicit activity requiring both cognitive and metacognitive processing. In order for the information system to adaptively intervene in learners' knowledge construction, the system must be able to capture their comprehension processes. The framework proposed in this paper is a promising approach toward this purpose.

In this study, we investigated the feasibility of the proposed systems through an initial operation verification by three learners. In the future, we plan to implement the learning materials in existing learning support systems, such as the one developed by Aburatani et al. (2019), to investigate learners' comprehension processes by analyzing learning logs. In addition, we need to consider the functions of an intelligent learning support system (Requirement 3) that provides feedback based on the learners' estimated comprehension processes and the state of knowledge.

References

Aburatani, T., Seta, K., Hayashi, Y., & Ikeda, M. (2019). Implementation of a System for Diagnosing Learning Behaviors of Thinking between the Lines Based on Presentation Design Tasks. *IEICE Transactions on Information and Systems*, Vol.J102-D, No.4, 359-363 (in Japanese).

Antonietti, A., Colombo, B., & Nuzzo, C. D. (2015). Metacognition in self-regulated multimedia learning: Integrating behavioural, psychophysiological, and introspective measures. *Learning, Media, and Technology*, 40, 187-209.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*, 219-243.

Hayashi, Y., Seta, K., & Ikeda, M. (2018). Development of a support system for measurement and analysis of thinking processes based on a metacognitive interpretation framework: a case study of dissolution of belief conflict thinking processes, *Research and Practice in Technology Enhanced Learning*, 13-21.

Kayashima, M., Inaba, A., & Mizoguchi, R. (2008). A Framework of Difficulty in Metacognitive Activity. *Transactions of Japanese Society for Information and Systems in Education*, 25(1), 19-31.

Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of picture labeling on science text processing and learning: Evidence from eye movements. *Reading Research Quarterly*, 48, 356-389.

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, 41, 172-187.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning, Cambridge University Press*, 43-71.

Mudrick, N. V., Azevedo, R., & Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying successful multimedia learning. *Computers in Human Behavior*, Vol.96, 223-234.

Muroya, D., Seta, K., & Hayashi, Y. (2021). Semantically Enhanced Historical Cartoons Promoting Historical Interpretation. *Information and Technology in Education and Learning*, 1(1), Reg-p002.

Ohno, T. (2002). What Can Be Learned From Eye Movement?: Understanding Higher Cognitive Processes From Eye Movement Analysis. *Cognitive studies: bulletin of the Japanese Cognitive Science Society*, Vol.9, No.4, 565-579 (in Japanese).

Roderer, T., & Roebers, C. M. (2014). Can you see me thinking (about my answers)? Using eye-tracking to illuminate developmental differences in monitoring and control skills and their relation to performance. *Metacognition and Learning*, 9, 1-23.