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

EXAIT: A Symbiotic Explanation Learning System

Brendan FLANAGAN[a*], Kyosuke TAKAMI[a], Kensuke TAKII[b], Yiling DAI[a], Rwitajit MAJUMDAR[a] & Hiroaki OGATA[a]
[a]*Academic Center for Computing and Media Studies, Kyoto University, Japan*
[b]*Graduate School of Informatics, Kyoto University, Japan*
*flanagan.brendanjohn.4n@kyoto-u.ac.jp

Abstract: Explainable artificial intelligence has been gaining much attention as systems increasingly make high-stakes recommendation and decisions automatically in a wide range of fields, including education. Meanwhile, research into self-explanation by students as a beneficial intervention to promote metacognitive skills has a long history of research. In this paper, we propose a learning system that aims to bridge understanding between the cyber and physical world by facilitating symbiotic explanation between the EXAIT system and students using the system. A co-evolution cycle of AI recommendation and the self-explanation by students of answer processes is proposed to increase the motivation and awareness of students, and at the same time enhance the effectiveness of the system.

Keywords: Explainable AI, self-explanation, symbiotic learning, recommendation

1. Introduction

In recent years, artificial intelligence has begun to revolutionize many fields, such as: medicine, finance, legal, and education, with decisions being made by models for prediction, estimation, and recommendation. As the use of models in education systems has been increasing, there are greater calls for not only the transparency and interpretation of inner workings of AI systems, but also technologies to explain complex models and the basis of results to end user stakeholders (Wang, Yang, Abdul, Lim, 2019). Much of the research into explainable AI (XAI) has focused on verifying the rationale behind results from such systems, however in education there are also additional benefits of explanations from learning systems that should be considered, such as: students learning from the explanations given by the system. Previous work into intelligent tutoring systems has shown that student's motivation in self-regulated learning with the system can be improved through prompt and feedback mechanisms leading to higher achievement (Duffy & Azevedo, 2015). Therefore, we argue that XAI in education should involve two main facets: explanation to increase student awareness of their course of study, and the explanation of recommendations for the purpose of fostering trust and motivation for continued use of the system.

While the explanation of system recommendations is gaining much attention, research into self-explanation from students in the context of education has a long history since Chi's seminal work on the topic (Chi et al., 1989). In their meta-analysis of 64 research reports on self-explanation and its effect on learning outcomes, Bisra, Liu, Nesbit, Salimi and Winne (2018) highlighted the potential of using self-explanation in a range of learning contexts as a beneficial intervention was suggested, however problems with instructor guided self-explanation was acknowledged. It was recommended that the effect of the intervention could possibly be improved by further investigation into system generated self-explanation scaffolds. Also, the analysis of self-explanations to inform AI driven learning systems, while mentioned to by Chi et al. (1989) is still a topic of ongoing enquiry. There is potential for learning systems to gain insight from detailed descriptions of answer processes that are input as a part of self-explanations by students. Such a system could learn a form of explaining the answer process of a question in Mathematics from high performing students, and then provide the learnt explanations as a scaffold from students that are struggling to perform the same task. Answer process analysis, such as pen stroke input time series analysis in Mathematics to find stuck points (Yoshitake, Flanagan, Ogata, 2020), coupled with self-explanation in the context of the answer process time series could uncover possible weaknesses in prerequisite knowledge for the problem at hand.

In this paper, we propose a system called EXAIT: Educational eXplainable AI Tools, which aims to combine these two aspects of explanation in education into a learning tool that can co-evolve in a symbiotic manner through learner's self-explanation and AI generated explanation. Firstly, the explanation of AI recommendations of possible learning paths from the view point of fostering trust and learner awareness. This then leads to students completing a recommended task, such as: a Mathematics question using a stylus pen to input the working out and final answer into the system for evaluation. The system then prompts the student to self-explain their answer to the task by replaying their answer process interactively, while annotating points in time to indicate what knowledge was applied to overcome sub-problems. Time series analysis is then applied to the self-explanation and answer process data to extract information, such as: backtracking or stuck points where the student had problems during the answer process that could indicate problems with dependent or related knowledge. The final goal of the system is to complete the symbiotic explanation cycle by informing the AI recommendation model from the self-explanation analysis.

2. Related Work

Previous work has proposed theoretical frameworks for symbiotic learning systems (Wu, Yang, Liao, Nian, 2021), where a learner learns from the system while the system also learns from the learner by employing reinforcement learning. However, the main purpose of such platforms and frameworks is to optimize the performance and efficacy of the system by finetuning a model to the behavior of the learners. In the present paper, we propose a symbiotic learning system focusing not only on adjusting the performance of the system to the learner, but also facilitating mutual understanding through the use of explanations by the learner and system to each other.

In the field of learning analytics, the iSTART tutoring system was proposed to tackle problems faced by students when paraphrasing during reading comprehension by supporting the task through the analysis of students' self-explanation in the context of the target text. A combination of NLP methods for self-explanation analysis and scaffolded instructions, and practice of self-explanation were implemented to encourage the development of required skills for the tasks. An automated evaluation algorithm gave students' scores on the quality of their self-explanations, with those that had high relevance of topics to the target text receiving high scores, and those that were off topic or short in length being assigned lower scores. It was found that the system could support self-explanation in different texts of various fields areas (Jackson, Guess, & McNamara, 2010). In the EXAIT system, the analysis of self-explanation will focus less on the automated evaluation of the content, and more on how the system can inform learner models on the strengths and weaknesses. It is intended that the analysis of self-explanation would influence the recommendation of learning paths of the student, thus closing the gap in the explanation co-evolution cycle.

3. EXAIT System

The EXAIT system proposed in this paper builds on the LEAF framework that has been developed to support the distribution of learning materials, collection and automated analysis of learning behavior logs in an open and standards-based approach (Flanagan & Ogata, 2018). The main components of the framework are: Moodle LMS which acts as a hub for accessing various courses; the BookRoll reading system for learning material and quiz exercise distribution; an LRS for collecting learning behavior logs from all of the components; and the LAView learning analytics dashboard to provide feedback to students, teachers and school administrators. This framework enables the real-time collection, analysis of learning behavior and feedback direct to stakeholders. An overview of the EXAIT concept is shown in Figure 1, with ⑤ LEAF platform being the foundation of the whole system. On top of this there is an abstracted model layer made up of ③ data-driven and ④ model-driven aspects. The model-driven aspect comprises of two main models: the knowledge model that represents the structure in the form of a knowledge graph of concepts that are learnt within a course, and the student model to monitor learning progress of individuals by analyzing data collected in the LEAF platform. The data-driven aspect comprises of an evidence model that guides the system based on the past performance of interventions. The interaction layer at the top of the diagram shows the symbiotic nature of the system while students are learning. There are two key aspects of the interaction: the system's ability to ① explain recommendations or decisions that have been made to the stakeholders, and reciprocal explanation by the students in the form of ② self-explanation of answers that they have submitted while using EXAIT. This reciprocal explanation between the system and students forms the basis of the symbiotic process that is the fundamental core of EXAIT.

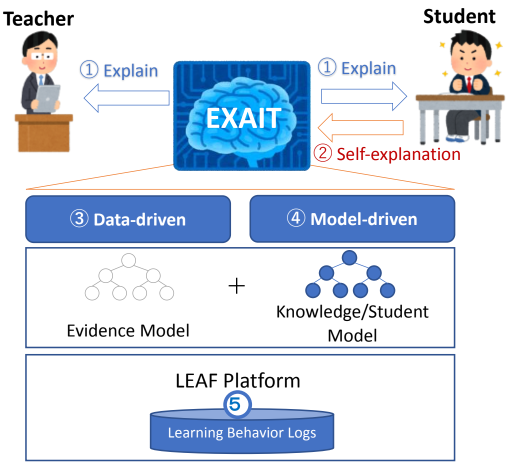

Figure 1. A High-level Overview of the EXAIT System.

The evaluation of the proposed system currently focuses on English and Mathematics education in Japanese secondary schools. The reason for this is that it is currently the target of the Japanese governments GIGA-school program to provide one computer per student to all children in compulsory education and mainly targeting secondary schools initially. Brod (2020) also highlights evidence from previous research showing that there is a favorable effectiveness of self-explanation when used in secondary school or higher. This confirms that the selection of secondary schools is a good fit with the proposed methods in the EXAIT system. In this paper, we will present the implementation of the EXAIT system for Mathematics education.

3.1 Recommendation Explanation in Mathematics

Exercise books used in the Mathematics course were uploaded to BookRoll and multiple-choice quiz questions were created to enable the collection of answers in learning log data. Students are also required to submit the working out that lead to the answer of the question as a hand written memo in the BookRoll interface as shown in Figure 2. The recommendation of learning paths in Mathematics has been popularized by the application of Bayesian Knowledge Tracing (BKT) in intelligent tutoring systems (Fischer et al., 2020), and is employed to model the degree of mastery of each mathematical skill in the EXAIT system based on the analysis of answers in learning log data. Question recommendations are made based on the probability that the student will correctly answer a question as determined by the BKT model, with extremely high or low probability questions having less weight in the recommendation.

We propose two main aspects from which recommendations should be explanation: Learner-centric and content-centric. The personal experiences of a learner can be extracted from learning logs to generate learner-centric explanations based on their past activity. An example of this could be that the learner has answered correctly questions on a skill that are prerequisites of the current recommended question. The EXAIT system collects evidence of previous cases of effective recommendation in the evidence model, which could also be used to explain to the learner about possible future learning path trajectories after completing the recommended question. Content-centric explanation can be generated by interpreting the parameters of the BKT model that has been trained, such as: the probability of transitioning from a state of unknown to mastery, the degree of forgetfulness or change that the answer might be a careless mistake. The clustering of these parameters will be analyzed to identify different types of questions for explanation, such as: the basic recall of knowledge in fundamental questions, to the application of knowledge in advanced questions. It is anticipated that this explanation will increase the learner's awareness of not only the type of question that is being recommended, but also the current phase of study.

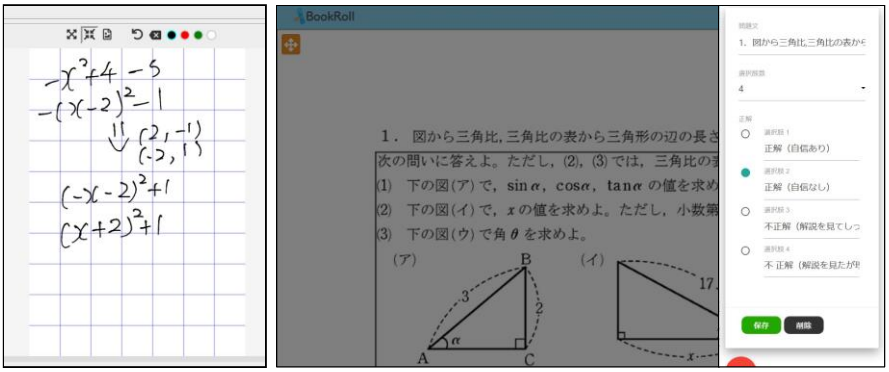

Figure 2. Handwritten Answer as a Memo (Left), Playback, and Self-explanation Input.

3.2 Answer Self-Explanation in Mathematics

As a part of the system, students are asked to explain their handwritten answers to questions that have been recommended by the system. The current user interface as shown in Figure 3 also includes the time delay analysis of pen strokes in the handwritten answer to indicate where students paused during the answer process (Yoshitake, Flanagan, Ogata, 2020). The student can playback and reflect on their answer by using the ① handwritten answer playback interface at the top. The rate of playback can be selected and users can also use the jump bar to skip to different parts of the answer process. When the playback is paused, the student has the option to create a new explanation for the particular step in the answer process. During playback the ② self-explanations of answers are displayed on the right of the screen and scroll through each of the explanations highlighting the current one as seen in Figure 3.

Previous research has shown that self-explanations can promote meta-cognitive skill use, as the students diagnose their answers, drawing on previously learnt knowledge that was applied to solve sub-problems (Chiu & Chi, 2014). It can also increase a student's awareness of learning by generating and completing self-explanation, using examples to convey new explanations of knowledge that has been acquired. The outcome of using examples to create self-explanations is that the student constructs more sophisticated knowledge through the process. However, while working with the class teachers, it became apparent that most students do not regard this task as worthwhile and it is viewed as more of a hassle. Also, students that have not grasped the required knowledge struggle with producing explanations when compared to high achieving students.

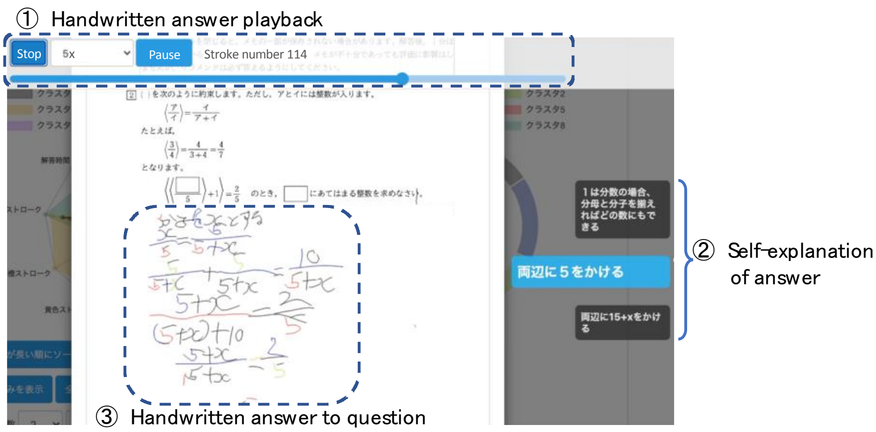

Figure 3. Handwritten Answer Analysis, Playback, and Self-Explanation Input.

Students that have not acquired sufficient knowledge to create self-explanations for their answer to a question may struggle or generate poor examples. Therefore, it may be necessary to support a student in the task, and could involve restricting the method of self-explanation as suggested by Wylie & Chi, (2014) ranging from open to limited respectively: Open-ended self-explanation (natural language); Focused self-explanation; Scaffolded self-explanation; Glossary/Resource-based self-explanation; Guided self-explanation. Currently, the EXAIT system has implemented an open-ended self-explanation interface with a high degree of freedom of expression. The self-explanations are input at points in time where the student deems appropriate without any prompting or guidance from the system. While this might be effective in a classroom where the instructor can guide the students through the process, there are potential problems with depending on such support as highlighted in previous research by Bisra, Liu, Nesbit, Salimi & Winne (2018). In the next iteration of the system, we plan to implement a guided process map based self-explanation interface that will provide students with predetermined keywords that are generated by the system. These keywords will be selected based on the required knowledge to complete the task. The student will then be prompted to arrange the keywords in order of the answer process and assign appropriate points in time where the knowledge was utilized.

4. Preliminary Study

In this section we describe two preliminary studies that were done to assess the use of recommendations provided the system without explanation. A recommender for both English and Mathematics was deployed to a secondary school in Japan. The students that participated in the study have past experience of using the LEAF system for at least the past academic year. The English reading recommender by Takii, Flanagan & Ogata (2021) was deployed as a book recommendation system for extensive reading in ESL classes. It is a knowledge map recommender method based on the vocabulary knowledge structure proposed by Flanagan et al. (2019). It was made available through two different dashboard systems: LAView which is a part of LEAF and GOAL a dashboard for Self Direction Skills (Majumdar et al., 2018). The Mathematics recommender proposed in previous sections was deployed only in LAView as GOAL had not previously been introduced to the target classes. Students were provided an orientation on how and why they should use the recommenders by the teacher of the course. As shown in Figure 4, initially there was high usage of the recommenders regardless of the subject or the system that was used for access. However, as time passed the usage of the recommender reduced even though the students were still engaged in studies. It should be noted that the time frame of the two studies is different in duration, with the English recommender being conducted over two weeks, and Mathematics over two months.

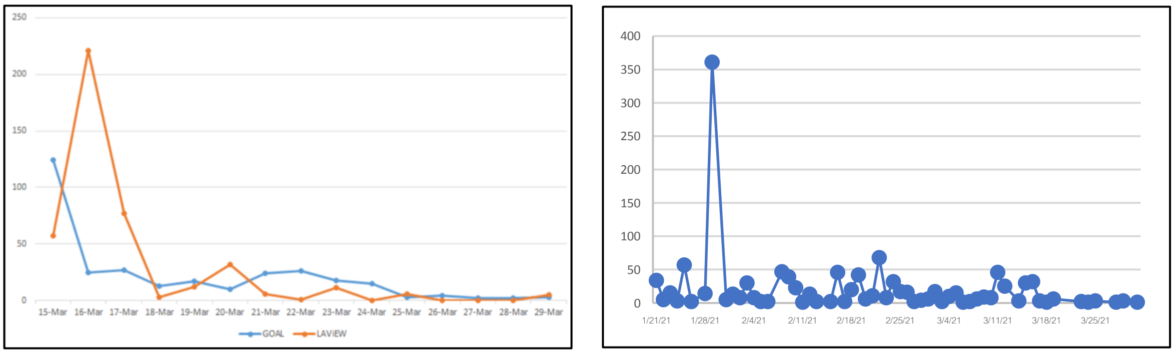

Figure 4. Recommender Usage (access frequency): extensive reading book recommendation for English class (left), question recommendation for Mathematics class (right).

5. Conclusion and Future Work

Recently, XAI has been gaining much attention in many different fields. In the context of education, we propose that there are some aspects of XAI that are unique to the field, such as the possibility of students learning from the explanations of a recommender. The implementation of AI system explanation and student self-explanation provide an opportunity to create a symbiotic learning system, which we proposed called EXAIT. In the present paper, we outline the current recommendation explanation and self-explanation components of EXAIT, and discuss possible future directions and challenges that remain. Two preliminary studies were conducted to see the use of recommendations in English and Mathematics courses without explaining the decisions to the students, and it was found that usage tapered off over time. In future work, we plan to formally evaluate learner-centric and content-centric methods of recommendation explanation. We also plan to implement and evaluate the effectiveness of computer-generated scaffolding for self-explanations by students.

Acknowledgements

This work was partly supported by JSPS Grant-in-Aid for Scientific Research (B) 20H01722, JSPS Grant-in-Aid for Scientific Research (Exploratory) 21K19824, JSPS Grant-in-Aid for Scientific Research (S) 16H06304 and NEDO JPNP20006 and JPNP18013.

References

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing Self-Explanation: a Meta-Analysis. *Educational Psychology Review*, *30*(3), 703-725.

Brod, G. (2020). Generative Learning: Which Strategies for What Age?. *Educational Psychology Review*, 1-24.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science*, *13*(2), 145-182.

Chi, M. T. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in instructional psychology*, *5*, 161-238.

Chiu, J. L., & Chi, M. T. (2014). Supporting self-explanation in the classroom. *Applying science of learning in education: Infusing psychological science into the curriculum*, 91-103.

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338-348.

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education*, *44*(1), 130-160.

Flanagan, B., Ogata, H. (2018). Learning analytics platform in higher education in Japan, *Knowledge Management & E-Learning: An International Journal*, *10(4)*, 469-484.

Flanagan, B., Chen, M. A., Lecailliez, L., Majumdar, R., Akçapinar, G., Ocheja, P., & Ogata, H. (2019). Automatic vocabulary study map generation by semantic context and learning material analysis. In *Proceedings of the 27th International Conference on Computers in Education (ICCE2019)* (pp. 698-702).

Jackson, G. T., Guess, R. H., & McNamara, D. S. (2010). Assessing cognitively complex strategy use in an untrained domain. *Topics in Cognitive Science, 2*, 127–137.

Majumdar R., Yang, Y.Y., Li, H., Akçapınar G., Flanagan B., Ogata H. (2018). GOAL: A System to Support Learner's Acquisition of Self Direction Skills, In *Proceedings of the 26th International Conference on Computers in Education (ICCE2018)*(pp. 406-415).

McNamara, D. S., Boonthum, C., Levinstein, I. B., & Millis, K. (2007). Evaluating self-explanations in iSTART: Comparing word-based and LSA algorithms. In T. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 227–241).

Takii, K., Flanagan, B., & Ogata, H. (2021). An English Picture-book Recommender System for Extensive Reading Using Vocabulary Knowledge Map. *Companion Proceedings of the 11th International Conference on Learning Analytics and Knowledge*.

Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019, May). Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-15).

Wu, J. Y., Yang, C. C., Liao, C. H., & Nian, M. W. (2021). Analytics 2.0 for Precision Education: An Integrative Theoretical Framework of the Human and Machine Symbiotic Learning. *Educational Technology & Society*, *24*(1), 267-279.

Wylie, R., & Chi, M. T. H. (2014). The Self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (2nd ed., pp. 413-432).

Yoshitake, D., Flanagan, B., & Ogata, H. (2020). Supporting Group Learning Using Pen Stroke Data Analytics. In *Proceedings of the 28th International Conference on Computers in Education Conference Proceedings (ICCE2020)* (pp. 634-639).