人工知能学会研究会資料
SIG-ALST-B503-07

Can LOD Based Question Generation Support Work in a Learning Environment for History Learning?

Corentin JOUAULT[1], Kazuhisa SETA[1, 2], and Yuki HAYASHI[2]

[1] Graduate School of Science, Osaka Prefecture University
[2] College of Sustainable System Sciences, Osaka Prefecture University

Abstract: The objective of this research is to build a system to support learners in self-directed learning of history in open learning space. The system supports learners with inquiry based learning using automatically generated questions. All the support provided by the system generated using the LOD. The evaluation of the system aims to verify the feasibility of LOD based support and its effectiveness. The results confirmed that LOD based support was feasible and that the question support improved engagement and historical deep thinking of the learners.

1 Introduction

When conducting self-directed learning in open learning space, one of the difficulties is that it requires learners to plan their learning in an unfamiliar domain. It requires learners to decide their learning objectives while they are learning. It is difficult for learners because they cannot easily judge the importance of the topics they encounter with their low knowledge of the domain. This problem is increased in open learning space because the quantity of information is much larger and the information is not organized with clear learning objectives. Learners can easily become overwhelmed and discouraged during learning because they cannot plan their learning. One way to eliminate the difficulty is prompting question generation and answering activities [Roth 96]. It contributes to lightening the difficulties of self-directing learning by lessening the planning activities that the learners need to perform. However, creating good inquiry questions requires an understanding of the domain and most learners cannot create good questions by themselves.

For this reason, we previously created a question generation function to create and adapt questions to learners. In a previous paper [Jouault 16], we demonstrated that it is possible to generate good quality history questions automatically in open learning space using Linked Open Data (LOD). A history professor judged that the quality of the automatically generated questions was on the same level as the ones generated by humans.

Another problem of self-directed learning is that learners must stay engaged in the learning task to have fruitful learning outcomes. To raise the engagement of learners while keeping them motivated, the support should be adapted to the learners' interests and orient them to important information without forcing them.

Since the quality of the questions was demonstrated to be high enough to support learners, these questions need to be provided to learners in an adequate way to support their learning. Thus, we created a learning environment: the Semantic Open Learning System (SOLS) to support learners in self-directed learning of history with adapted question support.

The research question to be answered in this paper is whether automatic question generation support is feasible and can help learners in self-directed learning of history in open learning space. We will first describe the system and how the question generation function was used. Then, we will describe an evaluation of the effects of the system on learners. The evaluation aims to verify two points: (1) the usability of the system that uses question generation using the LOD and (2) the learning effects of the question generation support on basic knowledge acquisition, historical considerations and engagement in history learning.

2 SOLS: Semantic Open Learning System

2.1 System overview

Fig. 1 shows a screen image of the learning environment. It is composed of four windows:

(a) **Document window:** It displays the learning material (Wikipedia document) the learner selected. All the

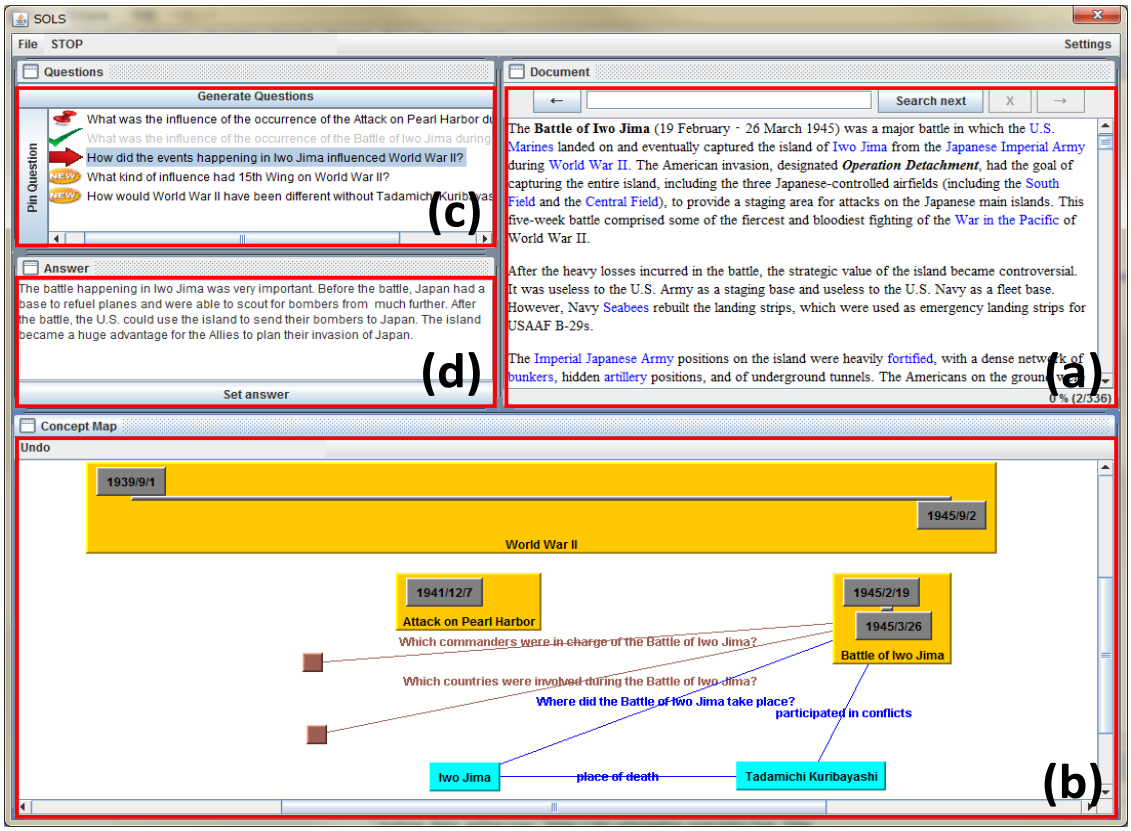

Fig. 1. System interface

links to a document about another concept appear in blue text. It provides usual Internet browser functions.

(b) **Concept map window:** Learners use this window to manage the concept map. The concepts in the middle are events. Events are represented on an automatically generated timeline built using data from the LOD. Other concepts are colored in blue and can be moved freely by the learner. The lines between two concepts are relation with the type of relation written at the center of the line. The concept map window also displays "shallow" questions that are designed to support basic knowledge acquisition.

(c) **Question window:** It displays a list of "deep" questions generated by the system designed to trigger deep historical thinking.

(d) **Answer window:** Learners use this window to answer their active questions selected in (c). Learners can write their answer to the question in natural language.

When using SOLS, learners can request question to support their learning when they want. The questions appear either in the concept map (b) or the question window (c) depending on the learner's request. By choosing one of the generated questions, learners actually decide a learning objective to be reached. Learners then browse documents in window (a) to find an answer to the question. Finally, learners answer the chosen question in window (d) or in the concept map (b). By repeating this process, the learners build their understanding by performing inquiry based learning in open learning space.

The purpose of the concept map building task for learners is to make them to describe their understanding explicitly. Research showed that building a concept map deepens the understanding of learners [Nesbit 06]. Building a chronology also reinforces learners' historical understanding [Stow 00].

In addition, the concept map is made fully machine understandable by using the LOD. Each concept added to the concept map has an ID that can be used to gather information on the LOD giving addition information about the concept to the system. Thus, it plays a key role to realize adaptive question generation that aims to deepen learners' historical considerations and help learners set learning objectives. The machine understandable concept map makes it possible for the system to assess the learners' knowledge and interests and adapt the generated questions accordingly.

2.2 Supporting learners with questions

All the questions are generated based on the LOD. The system can generate two sorts of questions depending on the learners' request: shallow questions in the concept map and deep questions in the question window.

First, on the learner's request, the system generates shallow questions appearing in the concept map to help the learners extend their basic knowledge. The generated questions will be displayed as a link between the concept and an empty node in the concept map. By displaying empty nodes, the system aims to trigger cognitive dissonances in the learner's mind. The empty node should make learners aware that they lack information and motivate them to fill the blank in their concept map. An example of concept map showing questions can be seen at the bottom of Fig. 1. These questions all require basic knowledge to be answered, such as a location for the questions "*Where did the Battle of Iwo Jima take place?*" The questions generated by this action are all questions about the concept selected by the learner. To answer the questions, the learners need to drag and drop the concept that answers it to the empty node and the system can verify the validity of the answer by using the data from the LOD.

Second, the question window (c) from Fig. 1 is designed to provide learners with adaptive advice generated based on the machine understandable concept map. Learners are provided with a list of deep questions in which they can select any that they consider interesting and try to answer it. In this situation, learners decide their learning objectives by choosing a question. The questions generated by the system were previously evaluated [Jouault 16] to be of a quality high enough to trigger historical thinking. The learners have access to good quality questions to direct their learning and their learning should be improved.

3 Experimental Setting

3.1 Objectives

The evaluation aims to verify four points:

1. The feasibility of LOD based support in a real learning scenario.
2. Whether the question support help learners develop their basic knowledge.
3. Whether the question support help learners develop their historical considerations.
4. Whether the question support can raise learners' engagement in history learning.

Concerning the objective (1), we aim to verify whether the functions implemented in the system work smoothly to not disturb learning and are considered useful by learners. The amount of LOD stored in the system, to enable support in open learning space, is around 80 GB. Thus, we need to confirm that the system can work smoothly without frustrating the learners. The evaluation is the first evaluation involving the system and many of the functions implemented in the system that use the LOD have not be proved to be effective before. For this reason, the feasibility of the system should be verified to prove whether the system can successfully support learners using only automatically generated support created using the LOD.

Concerning the objective (2), the hypothesis to verify is that learners using the question support can develop a stronger basic knowledge than learners using the system without advice. We expect learners using the system to have more developed context knowledge because the shallow questions trigger a cognitive dissonance for the learners, they become aware of gaps in their knowledge and each question gives them direction to find the missing information.

Concerning the objective (3), the hypothesis to verify is that learners using the question support should perform better in task that require understanding and not only basic knowledge such as essay writing requiring historical considerations. We expect the deep questions to prompt learners to develop their historical considerations during learning.

Concerning objective (4), the hypothesis to verify is that the question support improves the learners' engagement and interest in history learning. We expect the engagement of learners to be raised because the questions give objectives to learners to facilitate their learning and it should minimize the difficulties of self-directed learning.

3.2 Procedure

Table 1 shows the timetable of the experiment. This evaluation involved 24 university students. They were separated in two homogeneous groups of 12 learners each to form the control and experimental groups depending on their results to a basic knowledge test about World War II (WWII). Before starting the evaluation, the learners are instructed to learn about World War I (WWI) on Wikipedia to make them aware of their history learning skills.

During the experiment, learners are firstly explained about good historical considerations in history learning

and inquiry learning. Then, learners are taught how to use the different functions of the system they will use with a demonstration and practice on the system for the WWI topic.

Before learning with the system, learners are instructed that they will have to write a report about their historical considerations on WWII after learning. Both groups are informed about the report and are instructed to study with that objective in mind.

The learning task is to learn about WWII including learning about one topic among two in detail. The topics to learn in details are two important events from World War II: the Attack on Pearl Harbor and the Battle of Iwo Jima. We set a self-directed learning situation where learners can choose freely according to their interests. The learners have 90 minutes (equivalent to a usual classroom lesson time) to perform the learning activities using the system.

During learning, both groups use the system to build their concept maps and browse documents. The difference between the groups is the availability of the question advice. The learners in the experimental group can request questions at any time to support their learning. The learners in the control group do not have access to questions.

Table 1. Experiment timetable

Time	**Con. (n=12)**	**Exp. (n=12)**
Before experiment		
30 min	Study of WW I on standard browser	
10 min	Pre-Test about WWII for group formation	
Experiment		
5 min	Instructions: Good learning practices in history and inquiry learning	
35 min	• Demonstration of the system functions • Familiarization with the system	
	System **without** question advice	System **with** question advice
90 min	Learning with the system (Subject: WWII + Attack on Pearl Harbor or Battle of Iwo Jima)	
	System **without** question advice	System **with** question advice
20 min	• Questionnaire • Basic knowledge test (Subject: WWII and chosen topic)	
30 min	Report writing (Subject: Describe your historical considerations about WWII)	

At the end of the learning phase, the learners have to:

1. Fill a questionnaire about the system to evaluate their feeling about the different functions of the system using the 5-grade Likert scale.
2. Answer a high-school level test about their chosen topic and WWII to evaluate their basic knowledge.
3. Write a report about their historical considerations to evaluate whether they performed deep historical thinking activities.

4 Results

4.1 Feasibility

To evaluate the results concerning the usability of the system, both groups should be considered since both groups used the system to build their machine understandable concept maps.

The category "System usability" of Table 2 shows the results of the questionnaire about the usefulness of the system. The learners from both groups gave good average scores on the Likert scale to the answer to the questions concerning the usability and usefulness of the system. The learners in both groups benefited from using the system while the average marks of the learners in the experimental group using the question advice were higher than the learners in the control group. It means that the system functioned as intended and was able help learners.

The category "Concept map usefulness" of Table 1 also shows that learners judged that building the concept map was useful. Most particularly, the learners felt that the timeline created using the LOD was particularly useful. It suggests that they felt the concept map and timeline building is helpful to improve their understanding of the topic.

The category "Questions' quality" of Table 2 shows the results of the questionnaire for the questions only concerning the experimental group. It confirms the results of the evaluation of the questions which judged that the questions did not appear unnatural or nonsensical.

4.2 Effects on basic knowledge acquisition

The category "Shallow question usefulness" of Table 2 shows the results of the questionnaire concerning the shallow questions, which aim to improve learners' basic knowledge. Learners judged that the shallow questions were useful and motivated learners in developing their basic knowledge and extend their concept map by

Table 2. Questionnaire results

	Questionnaire item	Con.		Exp.	
		Avg.	**SD**	**Avg.**	**SD**
System usability	Do you think using the system helped you understand the topic?	3.92	0.67	4.25	0.62
	Do you feel that the system made it easier to learn?	3.50	1.09	3.92	0.79
Concept map usefulness	Do you think the timeline part of the concept map helped you understand the events?	4.25	0.75	4.50	0.67
	Do you think building a timeline including the context is a good way of learning history?	4.33	0.65	4.50	0.67
Interest in history	Are you interested in learning history?	3.00	1.28	3.58	0.90
	Do you want to learn more about history?	3.08	1.24	3.50	1.00
	Do you think using the system made you want to learn more?	3.08	1.24	3.67	0.65
	Do you feel using the system made you more interested in learning about history?	3.17	1.19	4.00	0.74
Interest in learning method	Do you think it is easier to perform inquiry based learning using the system?	3.33	0.78	4.25	0.62
	Do you think being aware of the questions is helpful to learn history?	3.67	0.78	4.58	0.51
Questions' quality	Do you think the questions seemed unnatural?	-	-	2.67	1.44
	Do you think the questions were nonsensical?	-	-	2.42	1.51
Shallow questions usefulness	Do you think the questions in the concept map made you realize what knowledge you needed to develop?	-	-	4.08	0.67
	Do you think the empty node that appears which each question in the concept map made you want to answer them?	-	-	4.00	0.74
	Do you think the questions not answered in free-text were useful?	-	-	3.75	0.75
Deep questions usefulness	Do you think the questions proposed by the system were about topic you were focusing on at the time?	-	-	3.42	0.79
	Do you think the questions answered in free-text were useful?	-	-	4.33	0.65

triggering cognitive dissonances. This confirms that the shallow question worked as intended.

However, even though the learners judged that the shallow questions were useful, the results of both the test about general knowledge of WWII and the topic chosen both show no significant difference between both groups for the tests about the chosen topic and about WWII. The question support in the short term use of the study does not seem to have a significant effect on the learning outcomes of the learners.

4.3 Effects on historical considerations

The category "Deep questions usefulness" of Table 2 shows the results of the questionnaire for the questions concerning the deep questions. The deep questions were judged useful by the learners. The learners also felt that the questions were adapted to their interests and knowledge as the questions were related to what they were studying. It means that the system was able to generate adapted deep questions to help the learners.

The reports were graded by a history professor with scores from 1 to 5. The results of the report show that there is little difference by the short term use between the grades of the reports from both groups (Control: 2.08, Experimental: 2.33).

More detailed analysis of the content performed by the history professor that graded the reports is shown in Table 3. The evaluator categorized the contents of the reports in 5 categories:

1. Personal feeling: the report describes the learners' personal feeling about the events.
2. Fact enumeration: the report is mostly a list of facts described without reflection by the learner.
3. Lesson learned: the report describes the lesson that should be learned from the events and makes the connection between the events and the current situation.
4. Historical considerations: the report describes the learner's historical considerations about the topic. The report contains the results of deep historical thinking from the learners.
5. Irrelevant: the report's contents cannot be categorized in another category or are confusing.

Some reports had contents that could be categorized in two different categories (3 for control group, 2 for experimental group). These reports are included in the count for both categories of the Table 3 in these cases. From 1 to 4 (excluding irrelevant content in 5), the history

Table 3. Categorization of reports content

	Personal feeling	Fact only	Lesson learned	Historical considerations	Irrelevant
Control (3 reports in 2 categories)	2	8	3	1	1
Experimental (2 reports in 2 categories)	0	3	2	6	3
Average number of questions answered (Experimental group only)	-	3.33	5.00	5.17	2.33

professor judged that the quality became higher from the viewpoint of historical considerations.

The detailed results show that even though the grades were similar, the content categories were different. Most learners in the control group (not using the question advice) wrote reports that are mostly enumerations of facts with little historical considerations. On the other side, many of the learners in the control group wrote reports containing deep historical considerations. The results also show that the more deep questions were answered by learners, the more their report is categorized in higher quality. It suggests that the deep questions prompt their historical considerations as intended.

4.4 Effects on engagement

The category "Interest in history" of Table 2 shows that learners in the experimental group showed more interest in history learning after using the system. The results show that the question generation support has a potential to raise engagement of the learners in history learning as the average mark for learners in the experimental group is higher. Since the learners in self-directed learning can choose to stop learning at any time, it is important to keep learners engaged in the learning task. Engaged learners will spend more time for learning and will develop their knowledge further.

The category "Interest in learning method" of Table 2 shows that using the system made the learners more aware of the usefulness of questions in self-directed learning of history. The results that the learners felt that the questions provided by the system helped them be aware of the importance of questions when conducting self-directed learning of history.

5 Concluding Remarks

The results of the evaluation showed that supporting learners in open learning space using support automatically generated using the LOD is feasible and can be useful for learners. Most learners judged that using the system was useful and that it helped them learn about history.

Even though the short term use of the question generation support did not have a significant effect on the learners' knowledge acquisition, it had an effect on the development of historical considerations of learners and motivated them to perform deep historical thinking to develop their opinion.

The question support also had an effect on the engagement of learners. Engagement has strong effects in self-directed learning because being engaged and interested in the topic lead learners to study for a longer time and learn about more topics

References

[1] Hmelo-Silver, C. E., Duncan, R. G. and Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, Vol.42, No.2, pp. 99-107.

[2] Jouault, C., Seta, K. and Hayashi, Y. (2016, to appear). Content-Dependent Question Generation using LOD for History Learning in Open Learning Space. *New Generation Computing*. Springer International Publishing.

[3] Nesbit, J. C. and Adesope, O. O. (2006) Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, Vol. 76, No.3, pp. 413-448.

[4] Roth, W. M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content, and student responses. *Journal of Research in Science Teaching*, Vol.33, No.7, pp. 709-736.

[5] Slotta, J. D. (2004). The web-based inquiry science environment (WISE): Scaffolding knowledge integration in the science classroom. In *Internet Environments for Science Education*, pp. 203-232.

[6] Stow, W. and Haydn, T. (2000). 7 Issues in the teaching of chronology. In *Issues in History Teaching*, p. 83. Routledge.

[7] Zimmerman, B. J. and Schunk, D. H. (Eds.). (2001). Self-Regulated Learning and Academic Achievement: Theoretical Perspectives. Routledge.