Diagnosis with Linked Open Data for Question Decomposition in Web-based Investigative Learning

Yoshiki Sato[1], Akihiro Kashihara[1], Shinobu Hasegawa[2], Koichi Ota[3] and Ryo Takaoka[4]

[1]The University of Electro Communications,[2] Japan Advanced Institute of Science and Technology, [3] Japan Institute of Lifelong Learning,[4] Yamaguchi University, Japan

[1]Yoshiki.sato@uec.ac.jp, [1]akihiro.kashihara@inf.uec.ac.jp, [2]hasegawa@jaist.ac.jp, [3]ko1ota@me.com, [4]ryo@yamaguchi-u.ac.jp

Abstract. In Web-based investigative learning, learners are expected to construct wider and deeper knowledge by navigating a great number and variety of Web resources/pages. On the other hand, they tend to search a limited number of them, which often results in limited knowledge construction. In order to make the investigation with an initial question elaborate, learners need to decompose the question into related ones. They also need to create a scenario like a table of contents implying the questions to be investigated and their sequence. We have built a model of Web-based investigative learning, and developed the system so far. However, it remains an open problem to diagnose learner-created scenario without preventing self-directed investigation. Toward this problem, this paper proposes a diagnosis method with Linked Open Data (LOD), and reports a case study whose purpose was to evaluate the diagnosis method.

Keywords: Web-based investigative learning, Linked Open Data, Self-directed learning, Diagnosis

1 Introduction

It is recently becoming more important to acquire the skill to utilize information, which is one of 21st-century skills [1]. Learning on the Web is particularly suitable for acquiring it [2]. The Web allows learners to investigate any question to learn from a great number and variety of Web resources in a self-directed way [3]. In Web-based investigative learning process, learners are expected to construct wider and deeper knowledge from their point of view [4], in which they select the Web resources/pages suitable for learning to navigate across them, and to integrate the contents learned at the navigated resources/pages by themselves.

On the other hand, learners tend to search a limited number of Web resources/pages for investigating a question, which often results in an insufficient investigation and limited knowledge construction. In order to make Web-based investigation with an initial question elaborate, it is necessary for learners to deepen and widen the question, which requires them to identify related questions to be further investigated during their navigation and knowledge construction [3]. This corresponds to decomposing the initial question into related ones as sub-questions.

M. Chang et al. (eds.), *Foundations and Trends in Smart Learning*,
Lecture Notes in Educational Technology, https://doi.org/10.1007/978-981-13-6908-7_15

In addition, learners are not provided with a scenario like a table of contents implying the questions to be investigated and their sequence. The learners accordingly need to create a scenario by themselves, which involves decomposing a question into the sub-questions for an elaborate investigation. Such a learner-created scenario would be helpful for learners to self-regulate their navigation and knowledge construction process [5]. But it is difficult for them to create their own scenario concurrently with navigation and knowledge construction.

In our previous work, we have proposed a model of Web-based investigative learning, and developed the system named interactive Learning Scenario Builder (iLSB for short) [6]. We have also confirmed iLSB could promote elaborate investigation and scenario creation [7]. On the other hand, it remains unclear whether scenario created by learners is appropriate.

This paper addresses a challenging issue how to diagnose the appropriateness of learner-created scenario in Web-based investigative learning. A general approach to this issue is to provide a correct scenario to compare with learner-created scenario. However, it is quite difficult to uniquely define it since Web-based investigative learning could bring about various question decomposition for an initial question. In addition, correct scenario provided would prevent learners from self-directed investigation.

Towards this problem, this paper proposes a method for diagnosing learner-created scenario with Linked Open Data (LOD), in which the appropriateness of relationships between a question and the sub-questions decomposed in the scenario is examined. We also report a case study whose purpose was to evaluate the validity of the diagnosis method. The results suggest that it could properly diagnose question decomposition.

2 Web-based Investigative Learning

First, we describe the model of Web-based investigative learning [6], and iLSB [7]. We then discuss the necessity for diagnosing learner-created scenario.

2.1 Model of Web-based Investigative Learning

This model includes three cyclic phases: (a) search for Web resources, (b) navigational learning, and (c) question decomposition. In phase (a), learners are expected to search and gather Web resources suitable for learning about an initial question using a search engine with a keyword representing it (called q-keyword), and to explore across their resources. In phase (b), they are expected to navigate the Web pages gathered in phase (a), and to extract keywords representing the contents learned in the pages to construct their knowledge. In phase (c), the learners are expected to find out some related sub-questions to be further investigated about the initial question, which are selected from the keywords extracted in phase (b). This corresponds to decomposing the initial question into sub-questions. Each sub-question is also investigated cyclically in the next phases (a) and (b).

The question decomposition results in a tree called question tree, which includes part-of relationships between the question and the sub-questions. The root of the tree represents the initial question. This tree also represents a learning scenario. Creating the tree corresponds to defining the initial question, which specifies what to investigate and how.

2.2 iLSB

We have developed iLSB as an add-on for Firefox. Fig. 1 shows the user interface of iLSB. iLSB provides learners with functions for scaffolding their investigative learning process: search engine, keyword repository, and question tree viewer.

Let us here describe how iLSB scaffolds question tree building with an example of investigation about “Global warming”. Learners are first expected to input “Global warming” as an initial q-keyword to the search engine, which is located in the root of question tree. iLSB allows them to search for “Global warming” and select/navigate the Web resources/pages. The learners second store keywords in the keyword repository, in which keywords represent the contents learned in the navigated pages. They then make inclusive relationships among stored keywords, and find out sub q-keywords to be further investigated. They also add the sub q-keywords to question tree viewer, and make part-of relationship from the root. The learners are next expected to investigate these sub q-keywords in the same way. As shown in Fig 1, the initial q-keyword “Global warming” is decomposed into three sub q-keywords such as “Greenhouse gas”, which is furthermore decomposed into the sub q-keyword “Carbon dioxide”.

2.3 Issue

In Web-based investigative learning, learners often investigate unrelated sub-questions for an initial question even if they use iLSB. This suggests the necessity of diagnosing learner-created scenario. A general approach to this issue is to provide a correct scenario, and to compare it with learner-created scenario. However, it is difficult to uniquely define the correct scenario since each learner can create his/her own scenario even for the same initial question. In addition, providing the correct scenario would prevent learners from self-directed investigation.

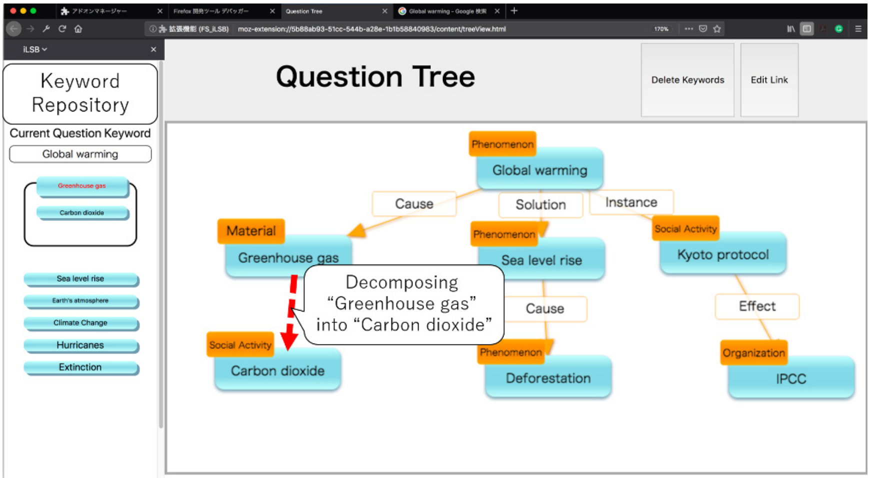

Fig 1. User Interface of interactive Learning Scenario Builder (iLSB)

Toward this issue, we aim to diagnose the appropriateness of question decomposition in learner-created scenario with LOD. The appropriateness of question decomposition is evaluated by means of DBpedia Japanese as LOD, in which iLSB calculates the relevance and similarity between the initial q-keyword and q-keywords decomposed, and the ones between the q-keywords and its parent q-keywords. The calculated results could be present to learners as feedback, which allows them to reflect on their question decomposition and scenario creation.

3 Method of diagnosis decomposed questions

Let us here describe how to diagnose learners' question decomposition with LOD.

3.1 Linked Open Data (LOD)

Linked Open Data is a set of structured data interlinking with related ones on the Web such as DBpedia Japanese [8] and YAGO [9]. In this work, we use DBpedia Japanese as LOD whose data are prepared for Japanese Wikipedia. Japanese Wikipedia is one of reliable [10] and structured resources for investigative learning. The data in DBpedia Japanese are expressed RDF (Resource Description Framework), which are encoded in a triple form of subject, predicate and object. A collection of triples can be represented as graph called RDF graph. In order to extract RDF data and operate RDF graph, it is necessary to send SPARQL query, which is a query language to operate RDF data stored in LOD. By means of SPARQL query, it is possible to measure distance between q-keywords in DBpedia Japanese, and to extract words related to q-keywords.

By means of distance between q-keywords, it is possible to calculate the relevance between q-keywords. In addition, it is possible to calculate the similarity between q-keywords by comparing related words of each q-keyword. The appropriateness of question decomposition is evaluated with the relevance and similarity between the initial q-keyword and q-keywords in question tree, and the ones between the questions and its parent question.

3.2 Framework of Diagnosis

Fig. 2 shows the framework of diagnosing the appropriateness of question decomposition.

The diagnosis is implemented as a function of iLSB. Let us explain the framework with an example of decomposing "Greenhouse gas" into "Carbon dioxide" when learners build a learning scenario about the initial question "Global warming" shown in Fig. 1. In order to evaluate the appropriateness of decomposing into the sub question "Carbon dioxide", in this case, iLSB calculates the relevance and similarity of part-of relationship between "Greenhouse gas" and "Carbon dioxide", and the ones between "Global warming" and "Carbon dioxide" by sending the SPARQL queries. The queries measure the distance and number of paths between these q-keywords, and obtain related words of each q-keyword. iLSB then diagnoses the appropriateness of question decomposition as one of three levels, such as appropriate, weak appropriate and unknown. iLSB gives the learners feedback about the appropriateness diagnosed.

4 Diagnosis of question decomposition

The appropriateness of question decomposition is evaluated with two criteria that are the relevance and similarity between q-keywords included in the question tree created by learners. Let us here describe how to calculate these criteria with DBpedia Japanese. We also describe a diagnosis procedure, which evaluates the appropriateness of question decomposition.

4.1 Relevance between q-keywords

Relevance is basically calculated by means of the distance and the number of paths in DBpedia Japanese between two q-keywords in question tree, which are obtained

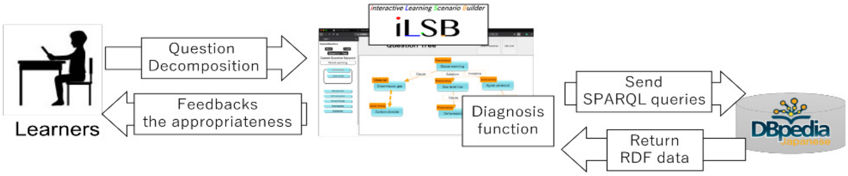

Fig 2. The framework of diagnosis

by SPARQL query. The query includes these keywords as its subject and object with any predicate. For example, the distance and number of paths between q-keywords "Greenhouse gas" and "Carbon dioxide" are measured with the following query1:

```
SELECT DISTINCT *
WHERE{
  {
    FILTER(contains(str(?subject_name),"Greenhouse gas"))
    FILTER(contains(str(?object_name),"Carbon dioxide"))
    ?subject rdfs:label ?subject_name.
    ?subject ?predicate ?object.
    ?object rdfs:label ? object_name.
   }UNION{
    FILTER(contains(str(?subject_name),"Greenhouse gas"))
    FILTER(contains(str(?object_name),"Carbon dioxide"))
    ?subject rdfs:label ?subject_name.
    ?subject ?predict1 ?middle.
    ?middle ?predict2 ?object.
    ?middle rdfs:label ?middle_name.
    ?object rdfs:label ?object_name.
  }
}
```

Query1: An example of SPARQL query

According to the distance and number of paths measured, iLSB decides the relevance as one of three levels: relevant, weak relevant, and unknown. We have conducted several preliminary examinations with several domains for deciding relevance level by means of the distance and the number of paths between keywords in DBpedia. The results suggest that keywords are relevant if the distance is 1, and they are weak relevant if the distance is 2 and the number of paths is more than 30. The relevance of keywords is suggested as unknown if the distance is more than 3 or the number of paths is less than 30. Following the lessons learned from the preliminary examinations, iLSB decides the relevance level of two q-keywords in question tree.

When the learners decompose "Greenhouse gas" into "Carbon dioxide" shown in Fig 1, for example, iLSB first sends the query1 to DBpedia Japanese. iLSB then obtains the distance and number of paths between these keywords. Since the distance between them is 1, in this case, the relevance between them is decided as relevant.

4.2 Similarity between Q-keywords

Similarity is defined with the intersection of two sets, each of which consists of related words for each of two q-keywords in question tree. The related words are obtained by morphological analysis of the results received with SPARQL query to DBpedia Japanese. It is evaluated by means of overlap coefficient as follows:

$$\text{overlap}(X, Y) = \frac{|X \cap Y|}{\min(|X|, |Y|)} \quad (1)$$

where X is a set of related words to one q-keyword, and Y is a set of related words to the other q-keyword. For example, the related words of “Greenhouse gas” are obtained with the following query2 and morphological analysis:

```
SELECT DISTINCT ?object
WHERE{
    FILTER(contains(str(?subject_name),"Greenhouse gas"))
    ?subject rdfs:label ?object_name.
    ?subject ?predicate ?object.
    FILTER(strStarts(str(?object),
     "http://ja.dbpedia.org/resource/"))
}
```

Query2: A SPARQL query for obtaining related words

According to the overlap coefficient calculated, iLSB decides the similarity as one of three levels: similar, weak similar, and unknown. We have also conducted several preliminary examinations with several domains for deciding similarity level by means of the value of overlap coefficient between keywords in DBpedia Japanese. The results suggest that keywords are similar if the value is more than 0.3, and they are weak similar if it is from 0.1. to 0.3. The similarity of keywords is also suggested as unknown if the value is less than 0.1. Following the lessons learned from these examinations, iLSB decides the similarity level of two q-keywords in question tree.

As for “Greenhouse gas” and “Carbon dioxide” in Fig1, for example, iLSB first sends the query2 and the one replacing the subject of the query2 with “Carbon dioxide”. iLSB then obtain words related to “Greenhouse gas” and “Carbon dioxide” from DBpedia Japanese and creates two sets including the related words with the morphological analysis. iLSB then calculates the overlap coefficient between the sets. Since the overlap coefficient between “Greenhouse gas” and “Carbon dioxide” is 0.5, in this case, the similarity is decided as similar.

4.3 Diagnosis Procedure

In diagnosing the appropriateness of question decomposition, we design a diagnosis procedure as shown in Fig.3. In this procedure, iLSB first calculates the relevance and similarity between question *i* (represented as q-keyword *i*) and the initial question (represented as the root q-keyword of question tree), and then calculates the ones between the question *i* and its parent question (represented as parent q-keyword). Depending on the calculated levels of relevance and similarity, iLSB decides the appropriateness of decomposing into the question *i* as one of three levels: appropriate, weak appropriate, and unknown. iLSB evaluates the appropriateness of each sub-question decomposed in the same way to diagnose question decomposition done by learners.

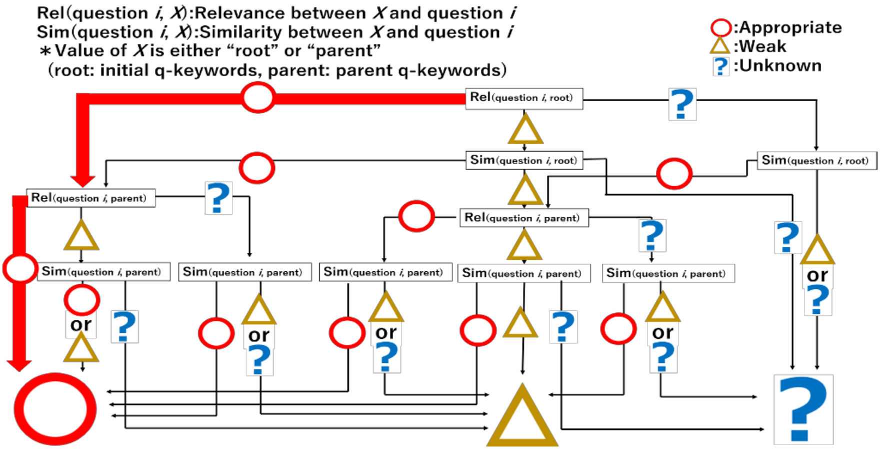

Fig 3. Diagnosis procedure

When learners decompose "Greenhouse gas" into "Carbon dioxide" in Fig 1, for example, iLSB calculates the relevance and similarity between "Global warming" and "Carbon dioxide", and the ones between "Greenhouse gas" and "Carbon dioxide". As shown in Fig.3, all these relevancies and similarities are evaluated as relevant and similar in this case. iLSB accordingly diagnoses the appropriateness of decomposing into "Carbon dioxide" as appropriate.

5 Case Study

5.1 Purposes and Procedure

In order to evaluate the validity of the designed procedure shown in Fig.3, we have had a case study where the appropriateness of question decomposition diagnosed with the procedure was compared to the one diagnosed manually. First, 9 graduate and undergraduate students in science and technology used iLSB to investigate two initial questions: "Tax" and "Judicial system". They then created their own learning scenarios for each question.

Second, these learner-created scenarios were diagnosed with the proposed procedure, in which each q-keyword decomposed in the question trees was examined. Third, three evaluators selected from the authors of this paper manually and carefully checked the appropriateness of question decomposition by referring to reliable Web resources, in which each q-keyword decomposed in the question trees was also examined. It was decided by majority of the three evaluators. In case each evaluator diagnosed it as different level, it was decided as weak appropriate.

Finally, we compared the appropriateness of the question decomposition diagnosed with the procedure and the one diagnosed manually, and evaluated the validity of the procedure. We used accuracy, recall, precision and F-measure as to the number of appropriateness level diagnosed for each sub q-keyword included in the 18 learner-created scenarios in total.

In evaluating the validity, in addition, we considered that the question decomposition diagnosed as weak appropriate was meaningful for learners to investigate in a self-directed way. In this work, we accordingly regard weak appropriate decomposition as appropriate one to evaluate the validity of the procedure. In other words, we compared the number of appropriate and weak appropriate decomposition with the one of unknown decomposition in learner-created scenarios.

5.2 Results and Discussions

Table 1 shows the numbers of appropriateness levels diagnosed with the designed procedure and manual diagnosis. The accuracy of question decomposition diagnosis with the diagnosis procedure and on manual was 77.8% (=224/288), which seems high. Table 2 also shows the recall, precision and F-measure as to the ratios of the diagnosis with the designed procedure to the manual diagnosis.

As shown in Table 2, the precision of appropriate question decomposition diagnosed with the procedure was almost 90%, and the F-measure was also quite high. The recall was also comparatively high. These results suggest that the question decomposition diagnosed as appropriate is evaluated validly.

As for question decomposition diagnosed as unknown, on the other hand, the precision was about 55%, which was lower than others, although the recall was about 75%. These results suggest that the designed procedure could properly diagnose unknown decomposition but misdiagnoses appropriate decomposition as unknown. The main reason is that the question decomposition includes keywords not to be obtained from DBpedia Japanese. This is the limit of LOD usage. We need to use other LODs, which is an important future work.

From the above discussion, the diagnosis results are not necessarily correct, but these seem instructive for learners to reflect on their question decomposition.

Table 1. Numbers of appropriateness levels diagnosed

<table>
<tr><td colspan="3" rowspan="3"></td><td colspan="3">Diagnosis procedure</td></tr>
<tr><td colspan="2">Appropriate</td><td rowspan="2">Unknown</td></tr>
<tr><td>Appropriate</td><td>Weak appropriate</td></tr>
<tr><td rowspan="3">Manual diagnosis</td><td rowspan="2">Appropriate</td><td>Appropriate</td><td>104(Q1:51,Q2:53)</td><td>30(Q1:6,Q2:24)</td><td>18(Q1:4,Q2:14)</td></tr>
<tr><td>Weak appropriate</td><td>23(Q1:18,Q2:3)</td><td>11(Q1:5,Q2:6)</td><td>27(Q1:6,Q2:21)</td></tr>
<tr><td colspan="2">Unknown</td><td>8(Q1:3,Q2:5)</td><td>104(Q1:51,Q2:53)</td><td>56(Q1:36,Q2:20)</td></tr>
</table>

Q1: Judicial system, Q2: Tax

Table 2. Recall, precision and F-measure of diagnosis procedure toward manual diagnosis

	Precision	Recall	F-measure
Appropriate & Weak appropriate	89.8%	78.9%	0.84
Unknown	55.4%	74.7%	0.64

6 Conclusion

This paper has addressed the issue how to diagnose the appropriateness of learner-created scenario without preventing learners from self-directed investigation with Web resources. We have also proposed a method for diagnosing learner-created scenario with LOD. In addition, we have reported the case study to evaluate the validity of the designed diagnosis procedure. These results suggest that the designed procedure could properly diagnose unknown decomposition but misdiagnoses appropriate decomposition as unknown. As a future work, we will ascertain whether the diagnosis method can promote reflection on question decomposition and self-directed investigation.

Acknowledgements. The work was supported in part by JSPS KAKENHI Grant Number 17H01992.

References

[1] Sarah, G. and Aman, Y: Computational Thinking and Media & Information Literacy: An Integrated Approach to Teaching Twenty-First Century Skills., TechTrends, Vol. 60, Issue 5, pp 510-516 (2016)

[2] Ananiadou, K. and M. Claro: 21st Century Skills and Competences for New Millennium Learners in OECD Countries: OECD Education Working Papers, No. 41, OECD Publishing, Paris (2009)

[3] Hill, J. R. and Hannafin, M. J.: "Cognitive Strategies and Learning from the World Wide Web", Educational Technology Research and Development, Vol. 45, No. 4, pp. 37–64 (1997)

[4] Henze, N. and Nejdl, W.: "Adaptation in Open Corpus Hypermedia", International J. of Artificial Intelligence in Education, Vol. 12, No. 4, pp. 325–350 (2001)

[5] Land, S. M.: "Cognitive Requirements for Learning Open-Ended Learning Environments", Educational Technology Research and Development, Vol. 48, No. 3, pp. 61–78 (2000)

[6] Kashihara, A. and Akiyama, N.: "Learner-Created Scenario for Investigative Learning with Web Resources", Proc. of the 16th International Conference on Artificial Intelligence in Education (AIED2013), LNAI, Vol. 7926, pp. 700–703 (2013)

[7] Kashihara, A. and Akiyama, N.: "Learning Scenario Creation for Promoting Investigative Learning on the Web", The Journal of Information and Systems in Education, Vol 15 Issue 1, pp. 62-72 (2017)

[8] DBpedia Japanese, http://ja.dbpedia.org/

[9] YAGO, https://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/

[10] Bhuminan, P, Christine, M, and Paul G. P et al.:"Continued value creation in crowdsourcing from creative process engagement", Journal of Services Marketing, Vol. 32 Issue: 1, pp.19-33 (2018)