

Generating Quizzes for History Learning Based on Wikipedia Articles

Yoshihiro Tamura¹, Yutaka Takase², Yuki Hayashi³,
and Yukiko I. Nakano²(✉)

¹ Graduate School of Science and Technology,
Seikei University, Musashino, Japan
dml26222@cc.seikei.ac.jp

² Department of Computer and Information Science,
Seikei University, Musashino, Japan

{yutaka-takase, y.nakano}@st.seikei.ac.jp

³ College of Sustainable System Sciences,
Osaka Prefecture University, Osaka, Japan
hayashi@kis.osakafu-u.ac.jp

Abstract. In intelligent tutoring systems (ITS), creating large amounts of educational content requires a large-scale and multi-domain knowledge base. However, most knowledge bases for ITSs are still manually developed. Aiming at reducing the cost of developing educational contents, this study proposes a method to generate multiple-choice history quizzes using Wikipedia articles. We also propose a method for assigning an importance measure to each relevant article based on hierarchical categories and the number of incoming links to the article. This is indispensable in generating quizzes that test basic knowledge of history. Finally, the results of evaluating these methods show that the proposed methods are useful in automatically creating quizzes for history exercise.

Keywords: ITS · Quiz generation · Wikipedia · History education

1 Introduction

In recent years, computers have become increasingly useful and popular in educational settings. Studies in intelligent tutoring systems (ITS) have attempted numerous approaches to estimate a learner's understanding and provide user-adapted, efficient learning environments. However, most knowledge bases for ITSs are still manually developed, and creating a large-scale, multi-domain knowledge base to generate educational content is costly. Therefore, in current educational systems, quizzes intended to test students' understanding of the subject matter are manually created, and their level of difficulty is manually determined.

With the aim of reducing the cost of developing education contents, this study proposes a method for generating history quizzes using Wikipedia articles. Wikipedia is a free online encyclopedia to which any users of the Internet can contribute. Since a large number of users contribute to Wikipedia, detailed information regarding a vast variety of topics is available on the site. It is customary for Internet users to refer to



Fig. 1. Introduction part of a Wikipedia article (Color figure online)

Wikipedia. Therefore, we think Wikipedia a rich knowledge base for automatically generating quizzes.

In this study, by focusing on Japanese history, we propose a method for creating quizzes where the answer to each prompt is the title of the Wikipedia article, using which the question was generated. To this end, we address the following issues in this paper:

1. Generating interrogative expressions from introduction parts in a Wikipedia article (text enclosed by the red rectangle in Fig. 1). We propose a method to transform a sentence in introduction part into an interrogative expression.
2. Generating answer choices for each quiz. In order to reduce the difficulty and ambiguity of questions, we propose a method to generate answer choices using Wikipedia’s tag information and triples in DBpedia, a project that aims to extract structured content from the information generated on Wikipedia.
3. Assigning an importance measure to each article based on the structure of the relevant category and link analysis. This is indispensable to generating quizzes that test basic knowledge of history and choosing appropriate quizzes for educational purposes.

Finally, we evaluate our proposed methods to show that they are useful for automatically creating history quizzes.

2 Related Work

As reviewed in [1], many attempts have been made in the area of question/problem generation in research on educational support tools. With regard to question generation for history education, Rus and Graesser [2] claimed that questions are categorized into two types: shallow questions (who, what, when, and where) and deep questions (why, how, what if). Since deep questions require descriptive answers, it is difficult for computer systems to judge their correctness. Thus, in this study, we address the problem of generating shallow questions regarding important people in history.

Limiting to shallow questions, this study aims to reduce the burden of manual educational content creation on educators. Previous studies on dialogue systems have contributed to this direction of research, and several methods have been proposed

for automatically generating questions using Web resources as knowledge bases. Misu et al. [3] proposed a dialogue system with a question-answer function and an information recommendation function based on Wikipedia resources. Higashinaka et al. [4] proposed a dialog system that generated quizzes asking for names of a person based on Wikipedia and online newspapers. Jouault and Seta [5] proposed a system that helped students with self-directed historical learning of history in an open learning space. In order to supply questions, they used semantic Web resources by combining DBpedia and Freebase contents. However, little research has addressed the generation of multiple-choice quizzes using Web resources in Japanese history.

Therefore, focusing on Wikipedia articles as a knowledge base, this study aims at automatic generation of multiple-choice quizzes in Japanese history domain. To accomplish this goal, we propose a method of quiz generation based on the results of morphological and syntactic analysis. In addition, in order to reduce the difficulty and the ambiguity of the generated questions, we propose a method for creating answer choices using RDB triples in DBpedia.

3 Wikipedia Database

In recent years, the correctness and reliability of Wikipedia articles has improved significantly. We thus used Wikipedia articles to develop our knowledge base for our history quizzes. To this end, we extracted Wikipedia articles on famous people from the Kamakura period, the Muromachi period, Nanboku-cho, the Sengoku period, the Azuchi–Momoyama period, the Edo period and the Bakumatsu, and the Meiji period. We also extracted articles of historical Japanese battles. Figure 2 shows the structure of our database, constructed using MySQL, and Table 1 lists the contents of the information stored in the database. The tables for “era”, “persons”, and “battles” are related to each other through relation tables. In this manner, we extracted 3,900 Wikipedia articles to register in the “person” tables and 500 articles for the “battle” tables.

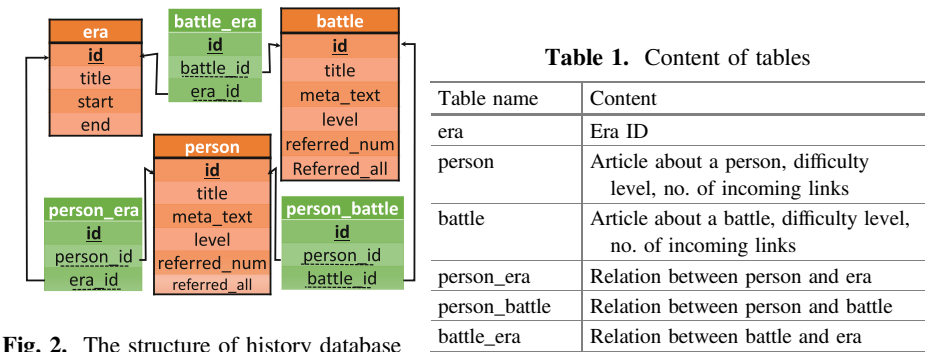


Fig. 2. The structure of history database

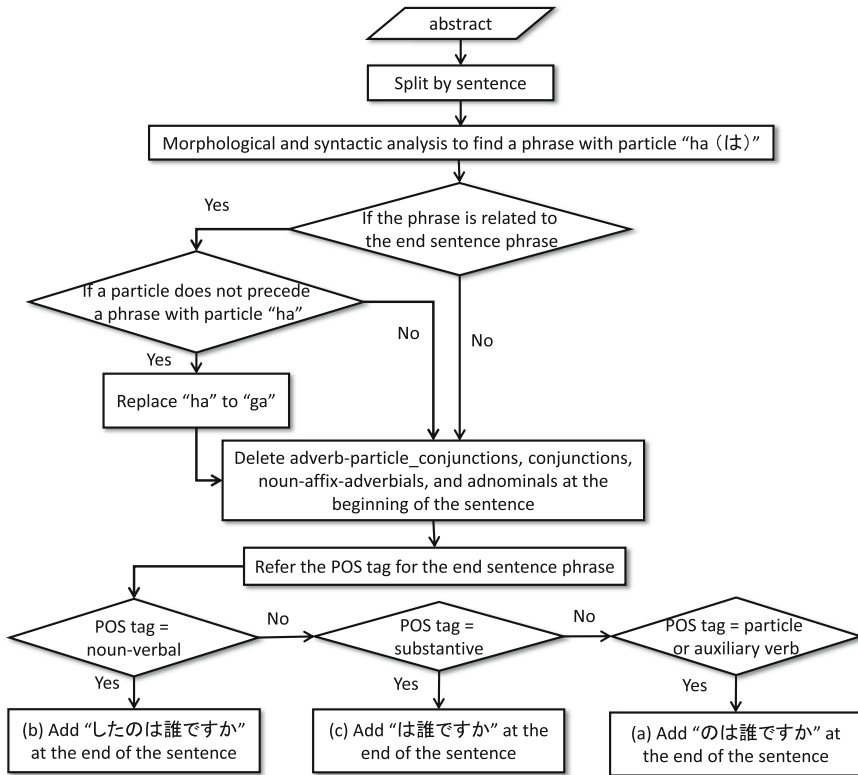


Fig. 3. Flowchart of quiz generation

4 Generating Quizzes

In this section, we detail our proposed method of generating quizzes from Wikipedia articles. The introduction parts of all Wikipedia articles are written in a very similar and consistent style. Moreover, the introduction part appears at the beginning of each page, because of which it is more likely to have been edited by many contributors, and thus is more likely to contain accurate information than other parts of the page. Based on this reasoning, we generated quizzes from the introduction part of articles.

We found two advantageous characteristics of introduction parts in generating quizzes:

1. Except for the first sentence, the subject of each subsequent sentence in the introduction text, which is the title of the article, is omitted.
2. Each sentence ends with either a noun or a noun phrase or with an assertive form, such as “-da” (だ), “-dearu” (である), etc.

The first observation implies that the first sentence of the introduction part of each article is suitable for generating quizzes, and the second observation suggests that

generic generation rules for interrogative sentences can be defined by using simple end-of-word transformations.

The introduction part contains sentences of the form “S (subject, name of the person) + V (a be-verb) + C (some explanation)”. By transforming these types of sentences, we generated a “who”-type quiz. The basic steps of the transformation are (1) deleting a noun phrase that includes the answer to the question, (2) adding a wh-interrogative expression to the end of the sentence, and (3) changing particles if necessary.

The flowchart of our quiz generation method is shown in Fig. 3. The introduction part is first divided into sentences. Morphological analysis and dependency (syntactic structure) analysis are then applied to each sentence. If a noun phrase including a postpositional particle 「は」 is found and, in the dependency analysis, relates to the end-sentence phrase, the postpositional particle of the noun phrase is replaced to 「が」. This is because two 「は」 particles should not occur in one sentence. In addition, conjunctions at the beginning of sentences (e.g. noun-affixed adverbials and adnominals) detected through morphological analysis are deleted. Following this, the end-sentence expression is changed in order to convert it into an appropriate wh-question. There are three cases:

- (a) if the sentence ends with a particle or an auxiliary verb, add the phrase “no (particle) + ha (particle) + dare (who) + desu (auxiliary verb) + ka (particle) [の是誰で`すか]” to the end of the sentence.
- (b) If the sentence ends with a noun-verbal, add a phrase “shita (verb) + no (particle) + ha (particle) + dare (who) + desu (auxiliary verb) + ka (particle) [したのは誰で`すか]” to the end of the sentence.
- (c) If the sentence ends with a substantive, add a phrase “ha (particle) + dare (who) + desu (auxiliary verb) + ka (particle) [は誰で`すか]” to the end of the sentence.

For instance, the first sentence of the introduction part of the article on Ieyasu Tokugawa, the founder and first shogun of the Togukawa shogunate, ends with “江戸幕府の初代征夷大將軍” (“the first shogun in Edo shogunate”). Since this is a noun phrase and ends with a common noun, “征夷大將軍 (shogun)”, transformation rule (c) (add “は誰で`すか?” to the end of the sentence) is applied, and the generated question is “江戸幕府の初代征夷大將軍は誰で`すか? (Who was the first shogun of the Edo shogunate?)”.

5 Choosing Important Articles for Quiz Generation

In order to generate quizzes that are appropriate as history exercises, it is necessary to select relevantly important Wikipedia articles. We assigned an article importance measure to each article using its number of referrals (incoming links) and hierarchical structure in Wikipedia.

5.1 Assigning Article Importance Using Wikipedia Hierarchical Structure

Each Wikipedia article has a list of categories that characterize its hierarchical relationships with other concepts. Therefore, these categories may be useful in estimating the importance of an article. However, while these categories are hierarchical, this structure is not entirely consistent because the categories of each article can be determined by any contributor. Because of that, we use “the main topic classification categories”, which are officially defined and are composed of nine basic categories, such as Discipline, Technology, Nature, Society, Culture, and History. In this study, we used the category “History” as the route category. “History of Japan” is in the fifth hierarchical level from the top level (Main topic classification → History → History by region → History of continent → History of Asia → History of Japan).

In the category of “History of Japan”, there is an article, “Japanese history”. We thus assumed that people mentioned in this article are the most important historical personages in Japanese history (level = 1). Since articles regarding eras, such as the Edo and the Kamakura eras, have direct links to the “Japanese history” article, historical personages, links to articles for whom are found within the article on the relevant historical era, are assigned the second-highest importance rating (level = 2). Other historical persons are assigned level 3.

5.2 Assigning Article Importance Based on the Number of Incoming Links

In addition to the category hierarchy in Wikipedia, the number of incoming links is used as an index of the importance of an article. These are links to the article in question from other articles in the database. We used the following two measures:

1. The number of tokens of incoming link
2. The number of types of incoming link

For example, if a person’s name, which is also a link to the relevant article, say “Tokugawa Ieyasu”, occurs two times in an article, the number of tokens of incoming link is 2 and the number of types is 1. In our history database, the number of tokens was registered in “referred_all” in the “person” table, and the number of types was registered in “referred_num”.

6 Generating Answer Choices for History Quizzes

A major problem in our quiz generation method was that the resulting quiz contained questions regarding small details that make the quizzes too difficult. More seriously, the quiz has ambiguity, so that there may be multiple answers. By having options from which to choose, it becomes possible for the user to select one correct answer even though the question might be ambiguous. Moreover, the level of difficulty of quizzes can be controlled by changing the choices, and different levels of quizzes can be generated using the same questions.

We used DBpedia Japanese to generate answer choices. DBpedia is a community project that aims to extract information from Wikipedia and publish it as linked open data (LOD). Information in LOD is registered based on the Resource Description Framework (RDF). The RDF describes the relationship between items of information in a resource using a triple: subject, predicate, and object. This study exploits these features in DBpedia to select answer choices.

Our algorithm for selecting answers choices is shown below:

- (1) Obtain a list of triples the subjects of which are matched with the title of a Wikipedia article.
- (2) Obtain a predicate–object pair from each triple and create a list of predicate–object pairs (hereafter called “PO”).
- (3) For each PO in the list, obtain a set of subjects whose predicates and objects are identical to those of a PO by searching the database.
- (4) For each PO, if either of the following conditions is satisfied, add the PO to a list of “selected” POs (hereafter caller “selected PO list”).
 - More than 50 % of the subjects obtained in (3) match the title of an article in the database.
 - More than 5 % of the article titles in the database match with subjects obtained in (3).
- (5) For all the combinations of two POs in the selected PO list, obtain the subjects using (3), and select those in common between the two POs.
- (6) If more than one and less than 51 subjects obtained in (6) match with the title of an article about a person in the database, save the subjects and the pair of POs as candidates for answer choices.

Note that Step (3) generates a set of persons who have something in common. However, it is possible that Step (3) generates a large number of candidates. Thus, we use Steps (4)–(6) to narrow down the set of candidates by computing the intersection of the results for all combinations of two POs. Using this algorithm, a set of highly similar answer choice candidates can be obtained. However, if the candidates are too similar, multiple answer choices can be correct. In order to resolve this problem, we implemented a process to delete candidates that were too similar.

7 Evaluation

7.1 Evaluation of the Questions

In order to assess the quality of our methods, we generated 100 quizzes containing questions asking for the names of historical persons and evaluated them in terms of grammatical correctness and educational effectiveness.

We randomly selected 100 articles from our database and applied our quiz generation algorithm to each. Since a quiz is generated from each sentence in the introduction part of a given article in our method, more than one quiz was generated if the

introduction part consisted of multiple sentences. We randomly selected 100 quizzes from among those generated, and asked the following questions about them:

- (i) Is the quiz grammatically correct? Is the sentence transformation completely successful?
- (ii) Is the quiz appropriate for learning? Is there any question with more than one correct answer choice? Is the content of the quiz educationally meaningful?

Three people attempted the 100 quizzes with regard to these two aspects. If a subject was unable to answer the quiz (because the quiz was too difficult), the quiz was determined to be educationally inappropriate. The results are shown in Table 2. The three reviewers, who were all native Japanese speakers, judged that 86.5 % of the quizzes were grammatically correct. This result suggests that our quiz generation method can generate grammatically correct quizzes. However, only 48 % of them were judged as appropriate as history exercises. Therefore, it is necessary to evaluate the content of the quizzes by professional reviewers of history and improve the method with regard to the educational appropriateness of the content.

Table 2. Evaluation of the quality of generated quizzes

	Evaluation aspects	Results
(i)	Grammatical correctness	86.5 %
(ii)	Educational appropriateness	48.0 %

7.2 Evaluating Article Importance

In order to evaluate the effectiveness of the proposed method in assigning article importance, we chose 100 historical persons using the following four methods, and compared our list with a list of historical persons who are mentioned in more than five history textbooks out of 11:

- (i) Using only the number of types of incoming links
- (ii) Using the levels of importance of articles in the category hierarchy and the number of types of incoming links. Under this condition, the articles are first sorted in order of category hierarchy, and articles at the same level in the hierarchy are then sorted based on the number of types of incoming links.
- (iii) Using only the number of tokens of incoming links
- (iv) Using the levels of importance of article in the category hierarchy and the number of tokens of incoming links. Under this condition, the articles are first sorted in order of category hierarchy, and articles at the same level in the hierarchy are then sorted based on the number of tokens of incoming links.

As shown in Table 3, 67 out of 100 historical persons selected by method (I) were found in the majority of Japanese history textbooks. On the contrary, when using the method (IV), 85 % of the selected historical persons were found in the majority of textbooks. These results suggest that our proposed method (method (IV)) can improve the accuracy of choosing important articles (historical persons) by 18 %.

Table 3. Evaluation of article importance

	Methods for assigning article importance	Result
(I)	referred_num	67 % (67/100)
(II)	referred_num + level	84 % (84/100)
(III)	referred_all	74 % (74/100)
(IV)	referred_all + level	85 % (85/100)

8 Conclusions and Future Directions

In this study, we proposed a quiz generation method using Wikipedia articles. Our method generates multiple-choice quizzes and thus generates answer choices as well as questions for quizzes. Experiments showed that our method can generate grammatically correct quizzes (more than 85 % of the generated quizzes were determined to be grammatically correct). However, more than half of the generated quizzes were determined to be educationally inappropriate. To improve the educational appropriateness of content generated by our method, we proposed a method to assign importance to articles using Wikipedia's hierarchy of categories as well as information regarding the number of incoming links for each article. The results showed that 85 % of the 100 most important historical persons ranked by our method were included in the majority of standard Japanese history textbooks. Furthermore, in order to reduce the difficulty and the ambiguity of the generated quizzes, we proposed a method to create answer choices.

In future work, we plan to integrate our quiz generation mechanism and the article importance estimation procedure into a system. We expect that by using important articles as source of quiz, the generated quizzes will be improved in the educational quality. Moreover, it is necessary to evaluate whether our multiple-choice quiz generation method can generate appropriate answer choices, and is effective in reducing the difficulty and ambiguity of the questions in the generated quizzes. With the aim of improving the answer choices, we will also develop a method for choosing words for distractors by using the categorical structure of Wikipedia and considering word similarities based on RDB triples in DBpedia.

Acknowledgements. This research was supported by CREST, JST.

References

1. Le, N.T., Kojiri, T.: Question and Problem Generation – State of The Art. In: Workshop Proceedings of the 18th International Conference on Computers in Education, pp. 47–55 (2010)
2. Rus, V., Graesser, A. (eds.) : The Question Generation Shared Task and Evaluation Challenge. ISBN: 978-0-615-27428-7 (2009)
3. Misu, T., Kawahara, T.: Speech-based interactive information guidance system using question-answering technique. In: Proceedings of IEEE-ICASSP 2007, pp. 145–148 (2007)

4. Higashinaka, R., Dohsaka, K., Isozaki, H.: Learning to rank definitions to generate quizzes for interactive information presentation. In: ACL 2007 Demo and Poster Sessions, pp. 117–120 (2007)
5. Jouault, C., Seta, K.: Content-dependent question generation for history learning in semantic open learning space. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) ITS 2014. LNCS, vol. 8474, pp. 300–305. Springer, Heidelberg (2014)