QA System Metis Based on Semantic Graph Matching at NTCIR 6

Minoru Harada† Yuhei Kato‡ Kazuaki Takehara‡ Masatsuna Kawamata‡
Kazunori Sugimura‡ Junichi Kawaguchi†
†Department of Integrated Information Technology, Faculty of Science and Engineering, Aoyama Gakuin University
‡Graduate School of Science and Engineering Master's Degree Course of Intelligence and Information, Aoyama Gakuin University

Abstract

We have developed Metis, a question-answering system that finds an answer by matching a question graph with the knowledge graphs. The question graph is obtained as a result of semantic analysis of a question sentence, the knowledge graphs are similarly analyzed from knowledge sentences retrieved from a database using keywords extracted from the question sentence. In retrieving such knowledge sentences, the system searches for and collects them using Lucene, a search engine, based on search keywords extracted from the question graph. To extract the answer, Metis calculates the degrees of similarity between the question and knowledge graphs to conduct precise matching. In this matching, the system calculates the degrees of similarity, which is the relative size of the similarity co-occurrence graph to the question graphs with respect to all combinations of nodes in the knowledge graph corresponding to those in the question graph. The system then chooses the knowledge graph with the highest degree of similarity and extracts from it the portion that corresponds to the given interrogative word. The system presents this portion as the answer.

Keywords: **Question answering, Graph matching, Semantic analysis, Semantic graph, Answer extraction**

1. Introduction

Recently, numerous studies have been in progress relating to question-answering systems, which extract answers out of an enormous set of sentences to answer a question sentence written in a natural language. The results from such research are announced at evaluative and other workshops, such as NTCIR's Question Answering Challenge (QAC) [8] and Cross Language Question Answering (CLQA) [9].

Though many methods have been announced so far, their basic concept is to search the Internet or newspapers for knowledge sentences, whose similarity in subjects suggests a relevant answer to the given question sentence. Then, those existing methods select the portions of the knowledge sentences thus discovered that correspond to the interrogative words of the question and present such portions as the answers.

Early research of this kind depended on the term frequency/inverse document frequency (TF/IDF) method in determining similarities between the question and knowledge sentences, which resulted in extremely poor precision of the answers provided. Kurata et al. [1] extracted an answer by obtaining the distance between nodes, which are sentence segments obtained as a result of the dependency parsing conducted after the extraction of answer candidates. In obtaining this node-to-node distance, they calculated a score for each answer candidate based on its distance from the search keywords extracted from the question sentence. Then, they extracted the answer out of the candidates in accordance with the scores obtained. However, the problem was that Kurata's distance calculation considered the modification relation alone in obtaining the distance and disregarded the role relation among the different nodes. For this reason, when a knowledge sentence had some redundant modification, the corresponding node-to-node distance became longer than it actually was, and the corresponding knowledge sentence ranked lower among the relevant sentences. As a result, Kurata's method was not always able to extract the answer correctly.

Murata et al. [2] conducted dependency parsing of a question sentence and a knowledge sentence extracted out of a database and then matched these two sentences in terms of the syntactic information to calculate the degree of similarity between the two. Murata obtained the answer based on the degrees of similarity thus obtained. This method obtained correspondences between the sentence segments of the two sentences matched and extracted as the answer, which was the segment that corresponded to the interrogative word. This method, however, considered the syntactic information only and disregarded the semantic relation between segments. Furthermore, the method skipped the semantic analysis of sentences and focused on the syntactic information alone as it obtained a degree of similarity between two corresponding segments. Thus, it was unable to obtain a semantic degree of similarity.

As described so far, processing a sentence in terms of its morphemes or syntactic information alone to extract an answer disallows the correct understanding of the semantic meaning of the sentence. As a result, methods employing such processing very often extract wrong answers, since they find correspondences in the words between the question sentence and knowledge sentence without considering the semantic similarity of words and the semantic relations among them.

2. Objective of our research

In our research to extract an answer, we developed Metis, a system that conducts a semantic analysis of a question sentence given in a natural language and that makes a full and precise matching of the semantic correspondence between the question sentence and a knowledge sentence.

In order to make a precise matching of a question sentence (e.g., "Who found out the plague bacillus and when?") and a knowledge sentence (e.g., "Shibasaburo Kitasato discovered the plague bacillus in 1894 in Hong Kong."), Metis conducts, with each of these two sentences, two conventional types of natural language processing, namely morphological analysis (we employed JUMAN from Nagao, Kurohashi et al. [10]) as well as dependency parsing (we employed KNP from Kurohashi, Kawahara et al. [10]). In addition, our system uses SAGE [4], developed by Harada's laboratory to conduct semantic and anaphoric analyses, whose results are output in the form of the semantic graph illustrated in the upper half of Fig. 1. In the semantic graph, each word is assigned a word meaning from the EDR computerized dictionary (a hexadecimal number of 6 digits) and the semantic relation (role) between two words is

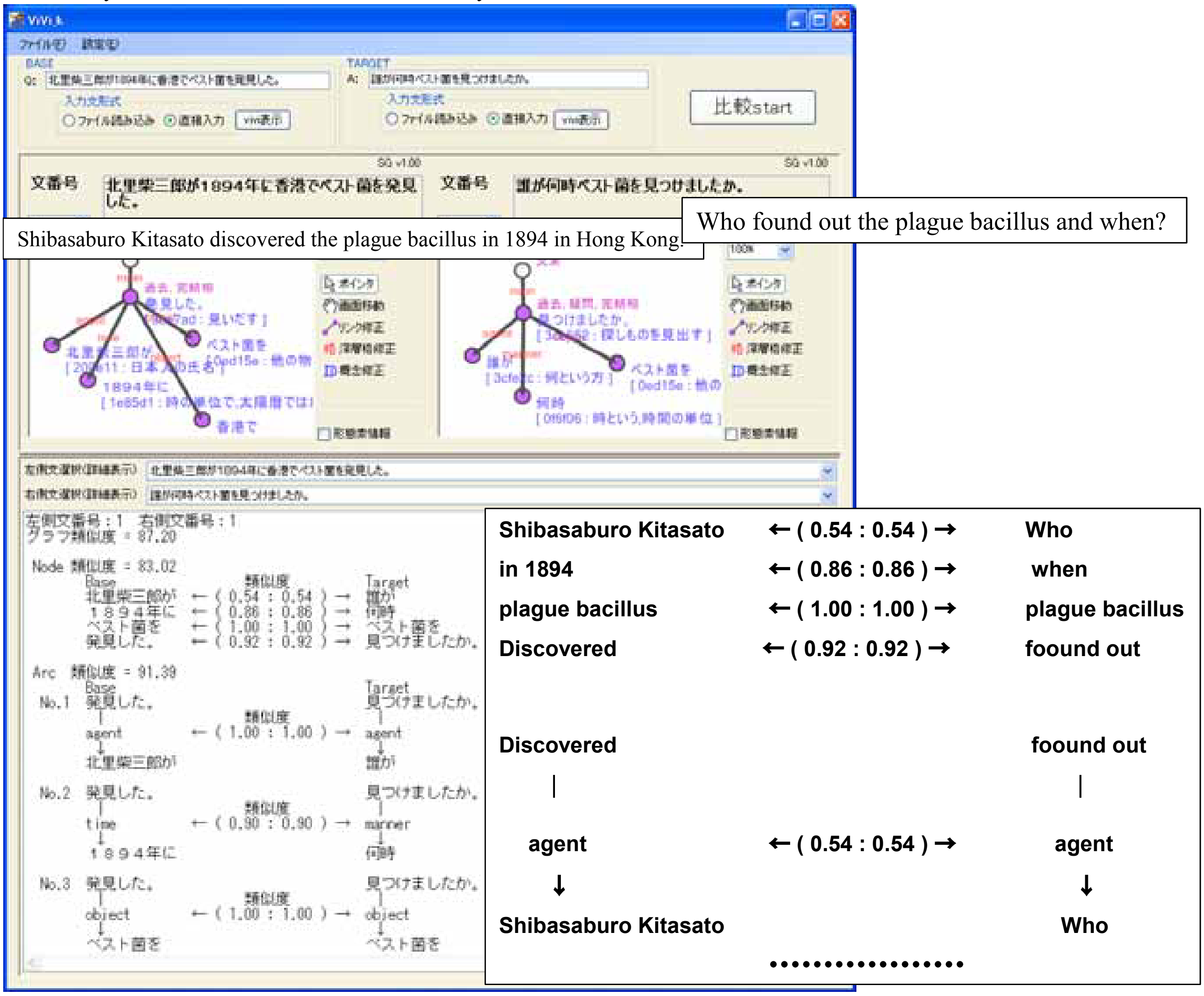

Fig. 1 Semantic correspondence between a question sentence and a knowledge sentence

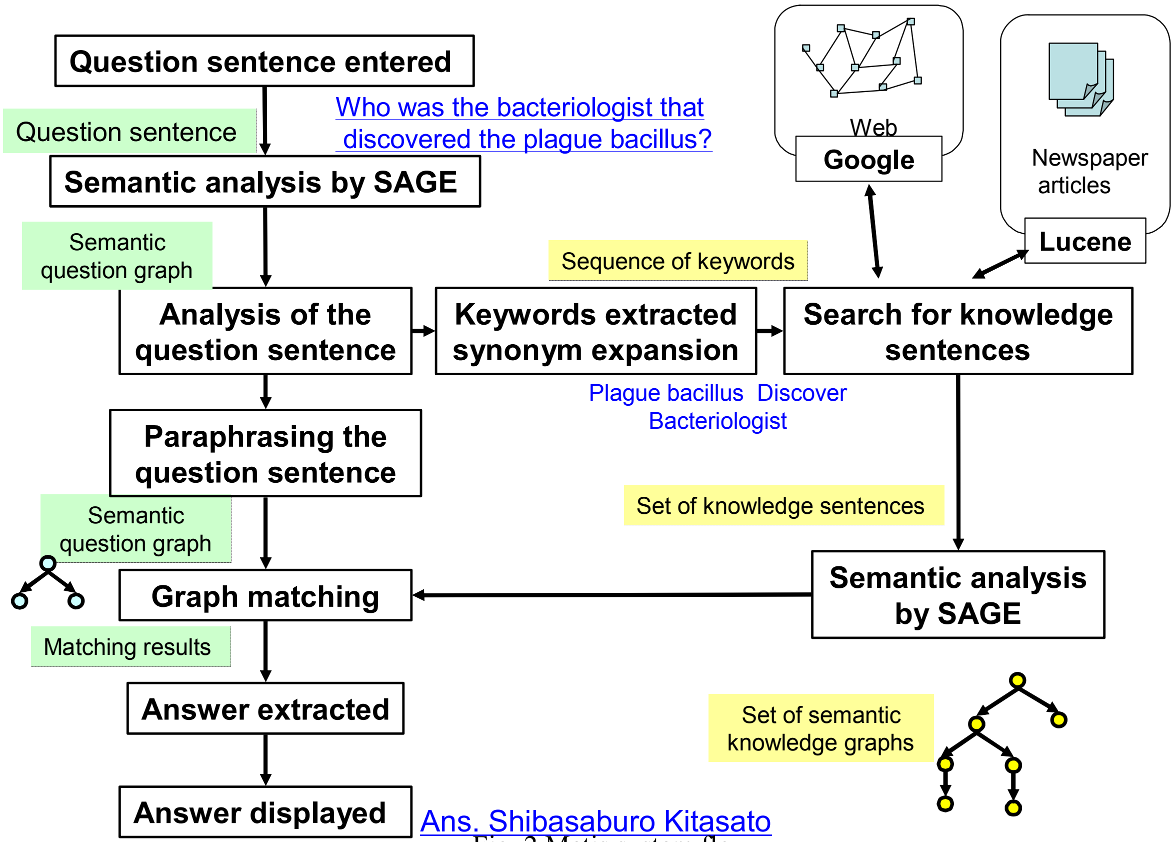

Fig. 2 Metis system flow

indicated by 30 or more deep cases. Metis measures the semantic similarity between a question sentence and a knowledge sentence in the form of the relative sizes of the common and similar portions of their respective corresponding semantic graphs against the question graph. For example, as shown in the lower half of Fig. 1, the common and similar portion graph for the two example sentences mentioned above consists of four node pairs (e.g., "Discovered ← (0.92 : 0.92) → found out?") and three arc pairs. The system calculates the word sense similarity of a node pair based on the distance to the common superordinate concept in the concept system tree of the EDR. (In this example, the distance is 0.92. The other 0.92 following the ":" denotes the distance reduced to reflect the mood difference. In this example, the two sentences are of the same mood and therefore the two distances are the same.) The system calculates the degree of similarity of an arc pair in accordance with the similar group of the deep cases. We define the particular pair as belonging to this group. After obtaining these degrees of similarity, Metis calculates their respective totals and then divides those totals by the number of nodes or the number of arcs within the question graph to obtain the node graph degree of similarity and the arc graph degree of similarity, respectively. The average of these two degrees is defined to be the graph degree of similarity. This way, Metis is capable of making the kind of judgment of the similarity between two sentences as humans, i.e., evaluating the similarity between corresponding words and inter-word relations in two sentences in terms of "Who did what, how, when, or where?"

3. Process flow of Metis

As the system flow chart of Fig. 2 shows, Metis first conducts a semantic analysis and extracts search keywords out of the question graph obtained from the semantic analysis. Then, using these keywords, the system searches the Web or newspaper data for relevant knowledge sentences. Next, the system conducts a similar semantic analysis of the knowledge sentences retrieved to produce a knowledge graph. Metis matches this knowledge graph with the question graph to extract the answer.

3.1 Classification of question types

In order to answer all types of questions, including those of the factoid type, which asks for a person's name, quantity, etc., as well as the why and how questions, and those asking for a definition, Metis classifies a question given at the beginning of its processing into one of several question types. As shown in Table 1, the system's classification consists of 12 types of factoid questions, as well as the why, how, and definition questions, each of which is treated as a single type.

Table 1 Classification of questions

Type of question	Example
Who	Who was the Italian physicist that invented the electric battery?
When	When does "The Little Match Girl" take place?
Where	Where is the capital of Indonesia?
What xxx is ---	What nationality was the company that acquired International Digital Communication?
Which xxx	Which state of the United States has the largest area?
What kind of	What kind of liquor is Beaujolais Nouveau?
What --- like	What is the shape of a "cube"?
What xxx ---	What team did they acquire?
How ---	How tall is Mt. Fuji?
How	How do you denote "Macao" in Portuguese?
How much	How much was Japan's current account surplus in 1998?
How many	How many people have a cellular phone?
Why	Why are you carrying an umbrella?
How	How did you get to the United States?
Definition	What is "K-1?"

3.2 Extraction of search keywords

For each of the nodes (sentence segments) in the question graph, Metis extracts search keywords. In this extraction, the system finds two types of keywords, "must" and "normal." A "must" keyword forms the core of the question and is always specified in searching the knowledge base. One or more "must" keywords are chosen from the question graph. For example, in the question, "What statues stand on both sides of Nandai Gate of Todai Temple in Nara Prefecture?," three "must" keywords are extracted: namely, "Nara Prefecture," "Nandai Gate of Todai Temple," and "statue." In addition, three "normal" keywords are extracted: namely, "both sides," "stand," and "statue." In this case, a search conducted with the normal keywords alone, "both sides," "stand," and "statue," could result in the retrieval of much unnecessary knowledge that is irrelevant to the knowledge asked for. Thus, in searching a knowledge base, we can reduce the quantity of irrelevant knowledge retrieved by including some "must" keywords among the search keywords used.

3.3 Acquiring knowledge sentences

When Metis searches its knowledge base, it chooses from two databases, one is the Web and the other is newspaper data. The search engine it employs for Web searches is Google [6] and the one it employs for a paper search is Lucene[5]. When creating indices with Lucene, the system conducts a semantic analysis of all the newspaper articles in advance, compiles them into a database, and then creates indices. The index keywords are created based on morphemes that are used for the semantic analysis. The keywords used in a database search are extracted on a morpheme basis that makes up the segment nodes in the graph. In this way, in the case of a compound word such as "the President of China," which is a segment node, Metis uses two search keywords, "China" and "President," and therefore is able to find a term such as "China's President." Also, in denoting a year in the Western calendar, people often write "82" in place of "1982." Thus, when a keyword is a Western year, Metis searches with two keywords, "1982" and "82."

Also, as it matches the graphs, Metis uses a concept system tree to obtain the degree of similarity between node words. For this reason, the system is able to determine that two different words with the same concept mean the same thing, as in the case of "discover" and "find out." Still, when it collects relevant knowledge sentences out of the database, the system conducts a search with the keyword "discover" and therefore is not able to find a sentence containing "find out." To solve this problem, Metis uses the EDR for synonyms to paraphrase keywords. Thus, the system conducts two searches, one using "discover" and the other using "find out."

By including synonyms among the search keywords, Metis is capable of collecting more knowledge sentences that seemingly contain the right answer.

3.4 Paraphrasing a question sentence

As Metis matches graphs, a higher degree of similarity with knowledge sentences might be obtained with a sentence of the "embedded" subject-predicate relation such as "Who did ---?" than with a sentence of the "presentational" relation such as "Who was it that did ---?" For example, suppose that the question is "Who was the bacteriologist that discovered the plague bacillus?" Then, the correct answer, "Shibasaburo Kitasato discovered the plague bacillus," has a different graph structure with respect to the modification and therefore is found to have a lower degree of similarity than it actually does. However, if we paraphrase the question into "Who discovered the plague bacillus?", the correct answer sentence has a higher degree of similarity. In order to prevent the lowering of the degree of similarity caused by a difference in the graph structure, Metis paraphrases question sentences. First, it roughly divides question sentences into two major types, "embedded" and "presentational," and then further classifies the "presentational" questions into two subcategories. As shown in Fig. 3, "presentational" questions are further classified into "complementary clause" and "adnominal clause," each of which is paraphrased as follows:

- In the semantic graph, a "complementary clause" question sentence is transformed from, for example, "Who is the person that invented the airplane?" into "Who invented the airplane?"

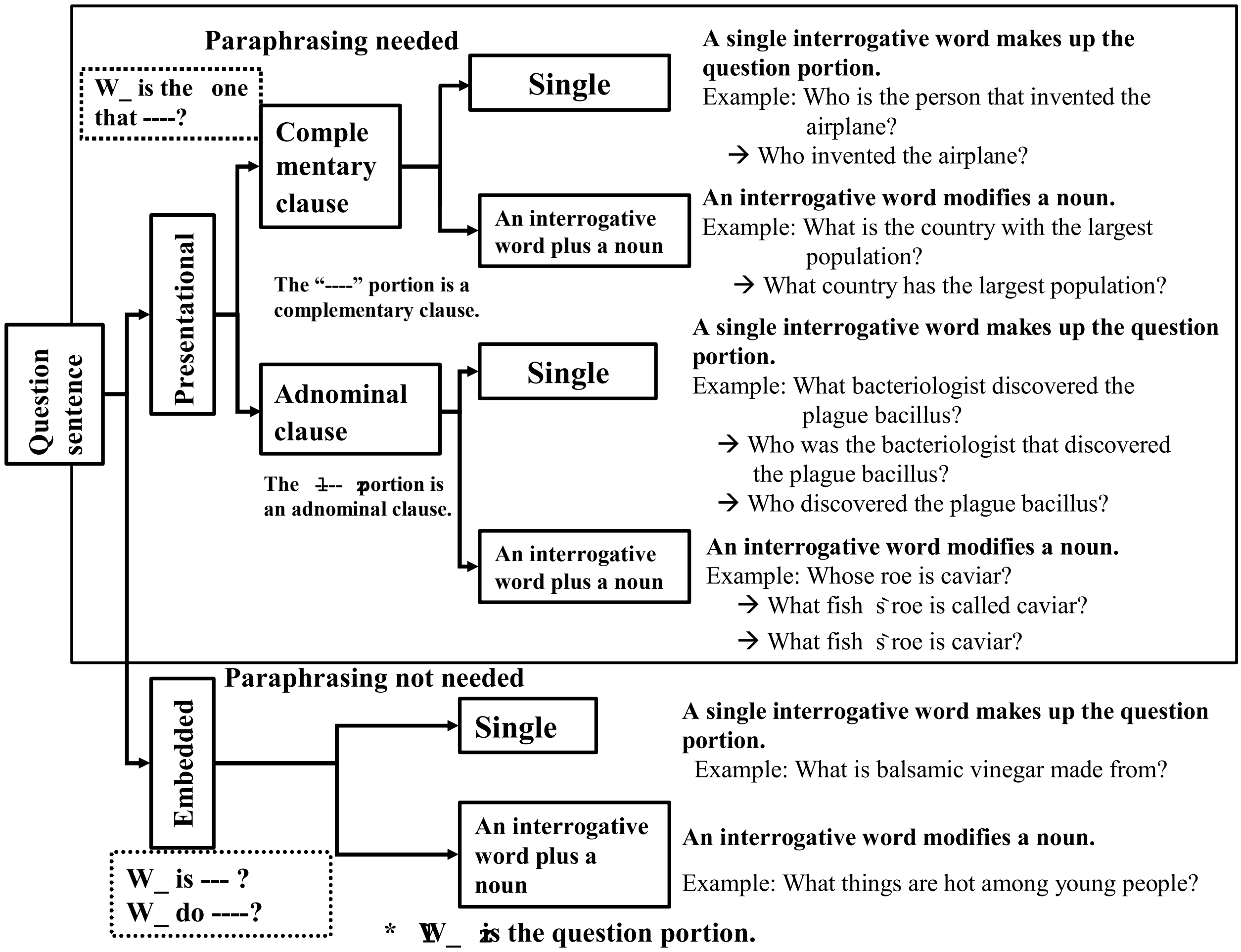

Fig. 3 Classification of question sentences to be paraphrased

- In the semantic graph, an "Who was the bacteriologist that discovered the plague bacillus?" into "Who discovered the plague bacillus?" or "What bacteriologist discovered the plague bacillus?"

Metis matches the transformed question graph as well as the one before the transformation with the knowledge graph to examine the degrees of similarity.

3.5 Matching of the question graph and the knowledge graph

In order to examine the similarity between the question graph and the knowledge graph obtained from the search, Metis matches these two graphs. First, the system calculates the degrees of conceptual similarity between the nodes in the question graph and those in the knowledge graph. Then Metis obtains the total of such degrees of conceptual similarity (called the "degree of node similarity") that exceed the threshold. The total is handled as the degree of node graph similarity. At the same time, the system obtains the total of the arc similarity degrees of the arc between the relevant nodes in the question graph and the arc between the knowledge graph nodes corresponding to both end nodes of the question graph arc. The total is handled as the degree of arc graph similarity. Then, after obtaining the degrees of the node and arc graph similarities, Metis obtains the sum of these two, which is called the degree of graph similarity. The degree of similarity between two nodes is defined as the conceptual similarity between the two nodes. (In handling proper nouns, the degree of similarity between two such nouns is based on their notations.) The degree of conceptual similarity between two concepts, C1 and C2, is obtained by the formula below, based on the two concepts' respective distance to the common superordinate concept, c (C1, C2), in the concept system tree of the EDR.

$$\text{Conceptual similarity} = \frac{2 \times d(c(c_1, c_2))}{d(c_1) + d(c_2)}$$

$d(c)$: Depth of concept c

The degree of arc similarity between two given arcs is specified in the right-hand column of Table 2, which lists the deep cases of similarity groups we defined. The degree of arc similarity is determined depending on which of these similarity groups the given two arcs' deep cases belong.

Table 2 Groups of deep cases

Name of group	Names of deep cases belonging to the group	Degree of arc similarity
Subject of an action agent	agent,o-agent,a-object, object,scene	0.90
Time sequence	time,time-from,time-to, duration, sequence, reverse, cooccurence, manner	0.90
Object of an action	object,goal,implement, material,source, o-agent, basis, beneficiary	0.85
Modifying expression	a-object,modifier,possessor, manner	0.90
Reason, cause	cause,reason, manner	0.80
Target of an action	goal,beneficiary,purpose, manner	0.85
Place	place,goal,from-to, location, scene,source, manner	0.90

Using the degrees of node and arc similarity, the equations below give the degree of graph similarity.

$$\text{Graph similarity} = \textit{Node} - \text{Graph similarity} + \text{Arc - Graph similarity}$$

$$\textit{Node} - \text{Graph similarity} = \frac{\sum(\textit{Node}\ \text{similarity} \times \text{Mood score})}{\text{Number of nodes of question graph}} \times 50$$

$$\text{Arc - Graph similarity} = \frac{\sum(\text{Arc similarity} \times \text{Mood score})}{\text{Number of nodes of question graph}} \times 50$$

The mood score is the score determined by comparing the moods of the nodes, such as “assertion,” “questions,” “past,” and so on. For example, consider the comparison of such nodes as “discovered,” “not yet discovered,” and “hoping to discover.” Since all of these nodes contain the word “discover,” they have high degrees of conceptual similarity among them. Still, “discovered” and “not yet discovered” actually mean opposite things. In a case such as this, applying a mood score of 1 or lower reduces the degrees of node similarity to bring the similarity recognized by the system closer to what humans feel about such sentences.

In the case of a factoid-type question, Metis uses the alternative concepts shown in Table 3 to calculate the degrees of conceptual similarity with such question nodes as “who,” “where,” and so on. For instance, if the type of question is “who,” the system employs the alternative concept of “person’s name” to better suit the question. This enables the matching of question and answer nodes in factoid-type questions.

Table 3 Alternative concepts for different types of questions

Type of question	Alternative concept given
Who	Name of a person, designation of a person, human
When	Time, point in time, quantity, unit of measurement
Where	Organization, name of a place, country
What xxx is ---	Concept of the node depended by the portion of question
Whitch xxx	Concepts of the head word and the sub-head word corresponding to the portion of question
What kind of	Particular thing, abstract thing, place, independent acting subject, state
What ---- like	Concept of the node depended by the portion of question
What xxx ---	Concept of the node depended by the portion of question
How ---	Quantity, unit of measurement, state
How	Event, thing, abstract thing, state
How much	Quantity, unit of measurement
How many	Quantity, unit of measurement, state

3.6 Matching of multiple sentences using anaphora

SAGE, the semantic analysis system employed in Metis, is capable of analyzing anaphora within a sentence. Using anaphora, our system conducts matching that involves the knowledge from multiple sentences.

The kinds of anaphora analyzed by SAGE include demonstrative pronouns and zero pronouns. For example, with respect to demonstrative pronouns, in the sentences, “Shibasaburo Kitasato was a bacteriologist. He discovered the plague bacillus,” SAGE adds the information that this “he” is “Shibasaburo Kitasato.”

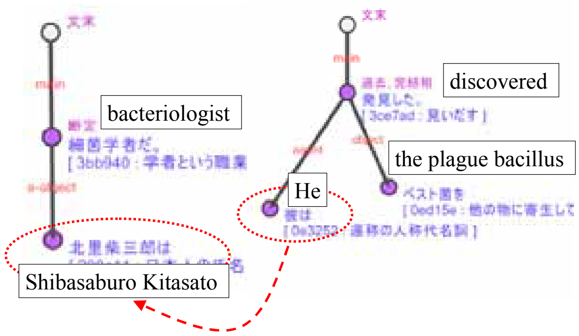

Fig. 4 Demonstrative pronouns

Using the same sentences as an example of a case with zero pronouns, “Shibasaburo Kitasato was a bacteriologist. He discovered the plague bacillus,” SAGE provides the complementary information that the subject of the action, “discovered the plague bacillus, z is Shibasaburo Kitasato. z Using information such as this in the matching of graphs, Metis is capable of extracting answers out of multiple sentences.

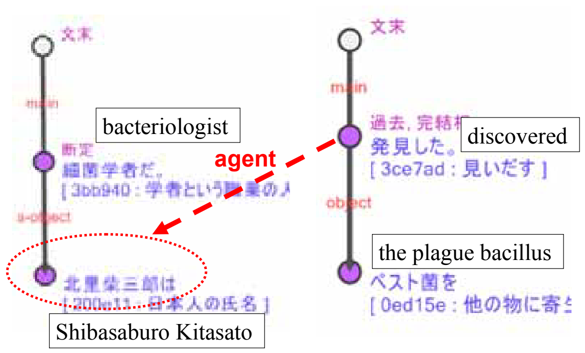

Fig. 5 Zero pronouns

3.7 Extraction of an answer

Using the results of the graph matching, Metis extracts an answer out of the knowledge graph with the highest degree of similarity with the question graph. As the basic principle in extracting an answer for a factoid type question, the knowledge node matched with the node of the question portion is extracted. For a question of the "why," "how," or "definition" type, Metis determines the subject topic node, which indicates the subject topic of the question within the question graph. Out of these knowledge nodes ("main knowledge nodes") matched with the subject topic node, the system determines, as the answer, the knowledge node ("ground node") that is connected to the subject topic node through the appropriate deep case. Now, the system extracts as the answer the subtree whose root is this ground node. The following section describes in detail how this answer extraction is conducted.

3.7.1 Extraction of an answer of the factoid type

In the case of a factoid-type question, as described earlier in the section of graph matching, Metis extracts as the answer a knowledge node matched with the question node, such as "who" or "where." In extracting an answer, the system extracts, together with the node chosen as the answer itself, the other node that modifies the answer (connects with it through a modifier case), if such a modifier node exists. For example, as shown in Fig. 6, regarding the

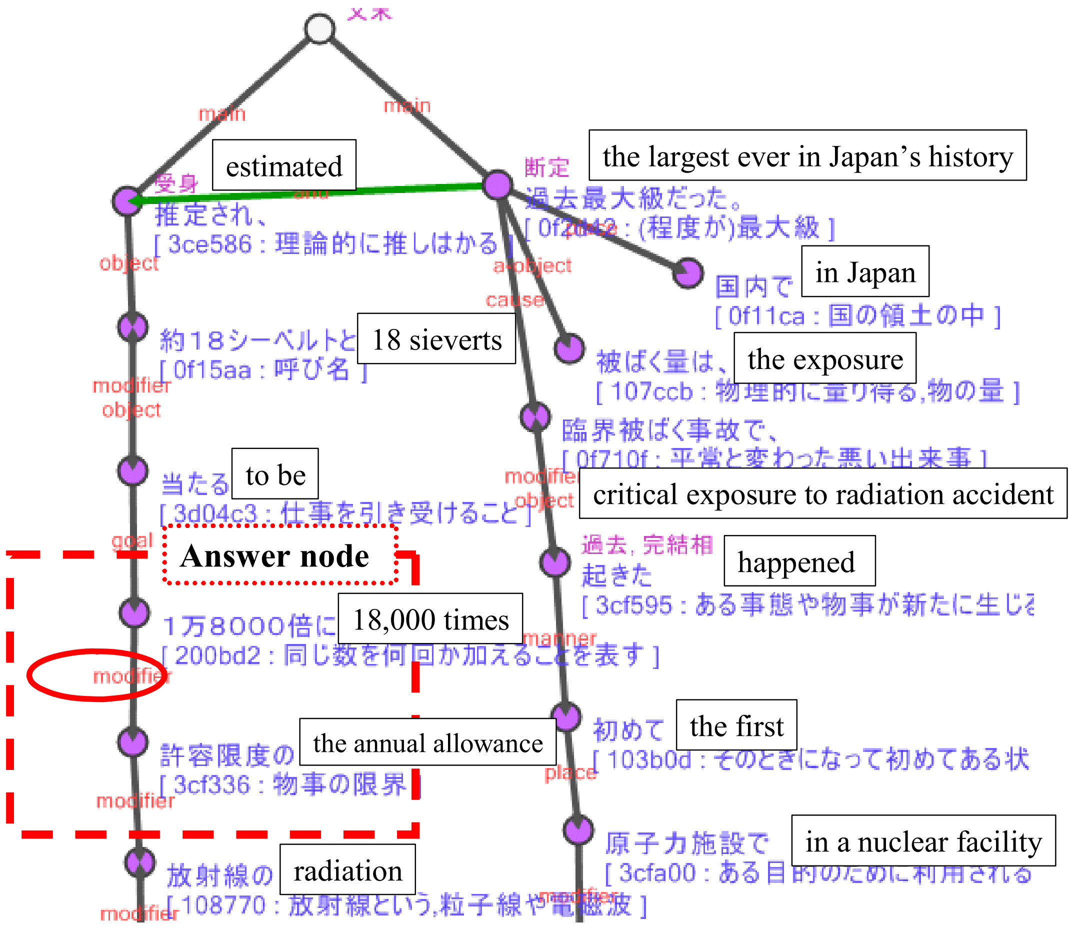

Fig. 6 Extraction of a factoid-type answer

question, “How many times more radiation was Mr. Hisashi Ouchi exposed to than the annual allowance, in the accident at the Tokaimura Uranium Processing Plant?”, suppose there is knowledge that “The accident was the first one in a nuclear facility in Japan of critical exposure to radiation, and the exposure dose is estimated to be some 18 sieverts, 18,000 times larger than the annual allowance of radiation for common people, the largest ever in Japan’s history of radiation accidents.” Then, the answer to extract should be “18,000 times.” In this case, since “larger than the annual allowance” modifies “18,000 times,” what is to be extracted is “18,000 times larger than the annual allowance.”

3.7.2 Extraction of an answer of the “why” type

In case of a question of the “why” type, Metis treats as the subject topic node the main predicative node in the question graph (which is one of the search keywords). Then the system extracts the main knowledge node from the matched knowledge graph that is matched with the subject topic node and the set of the subtrees within the knowledge graph whose root is the ground node connected to the main knowledge node through deep cases denoting a reason, namely, “reason,” “cause,” “manner,” “sequence,” “location,” and “sequence.” Fig. 7 shows an example of this. To the question, “Why are India and Pakistan opposed to each other?”, we have the knowledge sentence that “While Pakistan is the only nation in the world that approves of the Taliban government, India is close to the anti-Taliban coalition (the Northern Alliance), and India and Pakistan are slightly opposed to each other over the Afghanistan situation.” With the main knowledge node being “opposed to each other,” the system extracts as the answer the subtree whose root is the ground node connected to the main knowledge node through the “sequence” and “cause” cases. Thus, the answer extracted includes “While Pakistan is the only nation in the world that approves of the Taliban government, India is close to the anti-Taliban coalition (the Northern Alliance) ------- over the Afghanistan situation.” Also, in the case that no single knowledge graph contains a deep case denoting a reason, Metis extracts the sentence that follows the main knowledge node, if the SAGE context analysis shows that the relation between the main knowledge node and the sentence that follows is a “reason.” In this example, if the knowledge sentence is “India and Pakistan are opposed to each other over the Afghanistan situation, because Pakistan is the only nation in the world that approves of the Taliban government and India is close to the anti-Taliban

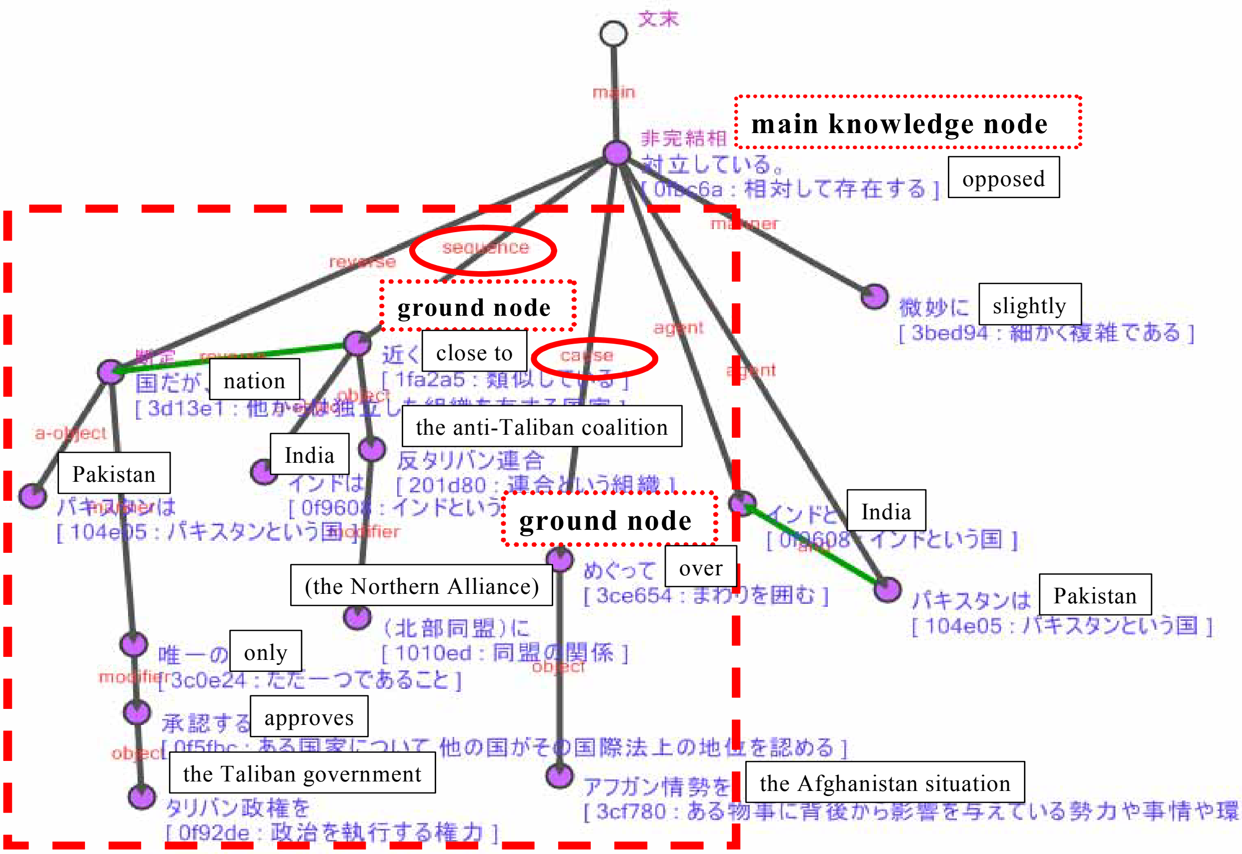

Fig. 7 Extraction of a “why” type answer

coalition (the Northern Alliance)", then "India and Pakistan are opposed to each other over the Afghanistan situation" is connected to the sentence that follows, "because Pakistan is the only nation in the world that approves the Taliban government and India is close to the anti-Taliban coalition (the Northern Alliance)" through the inter-sentence deep case "reason." Thus, the latter sentence is extracted as the answer.

3.7.3 Extraction of an answer of the "how" type

In the case of a question of the "how" type, Metis treats the main predicative node in the question graph as the subject topic node and extracts the set of subtrees within the knowledge graph whose roots are ground nodes connected to the main knowledge node through deep cases denoting a manner or method such as "implement," "sequence," "condition," "manner" and "scene." Fig. 8 provides an example of this. To the question, "How is dioxin generated?", we have knowledge that "Since people actually began to worry over dioxin generation in food wrap when it is microwaved ---." In this case, the extracted answer is "generation in food wrap when it is microwaved." It consists of the subtrees whose root is the ground node connected to the main knowledge node, "is generated," through the "condition" and "scene" cases.

3.7.4 Extraction of an answer of the "definition" type

In the case of a question of the "definition" type, Metis treats the word whose definition is being asked for as the subject topic node and extracts the subtree within the knowledge graph whose root is the ground node connected to the main knowledge node through the "modifier" case. Fig. 9 shows an example of this. The question, "What kind of a sport is skeleton?", asks for the definition of the word "skeleton." We have a sentence, "On the 10th, the All Japan Championship of 'skeleton,' a sport in which a competitor rides on a sled in a prone position and slides down a course of ice head-first, ---." Then, the answer to extract is "a sport in which a competitor rides on a sled in a prone position and slides down a course of ice head-first," since this is the subtree whose root is the "modifier" case of "skeleton." Also, in the case that there is a word connected to the main knowledge node through "a-object," that word is treated as the ground node and the corresponding subtree is extracted. In this case, in response to the

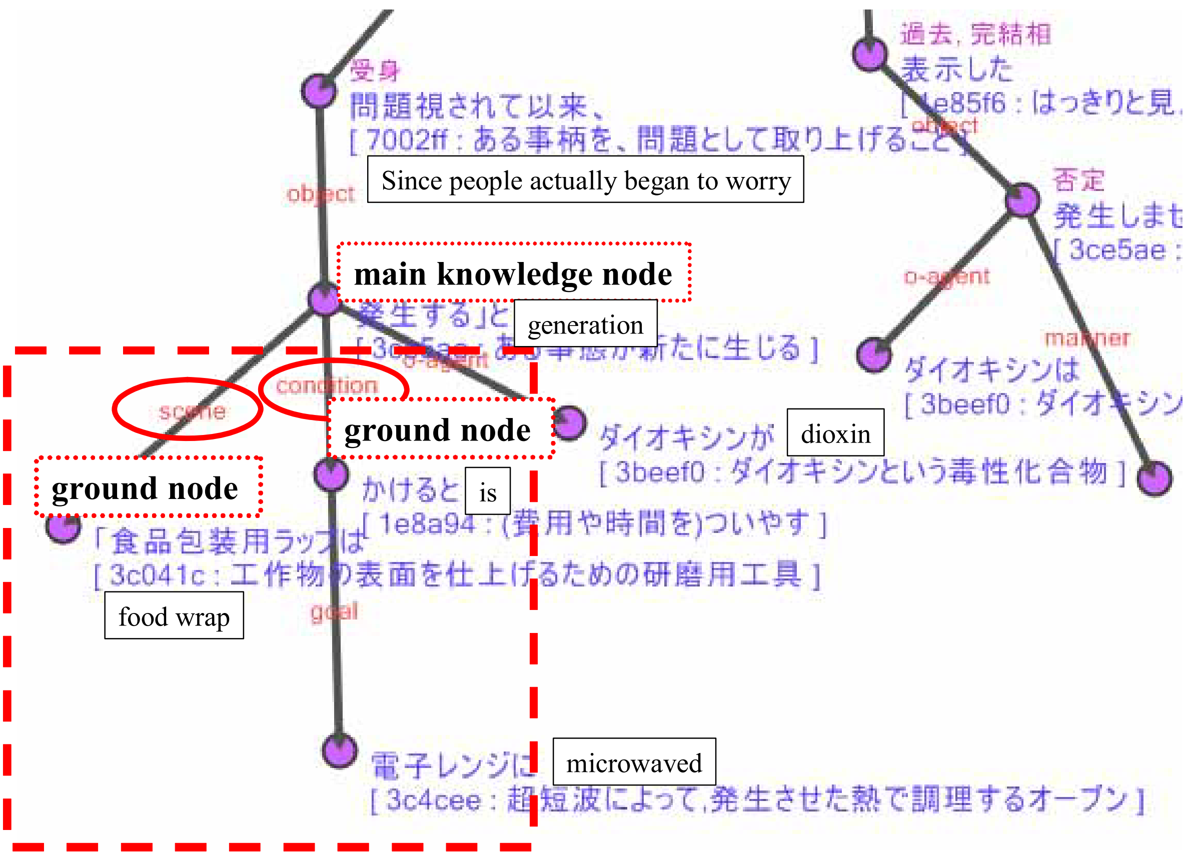

Fig. 8 Extraction of a "how" type answer

knowledge that "Skeleton is a sport in which a competitor rides on a sled in a prone position and slides down a course of ice head-first," the answer extracted is "a sport in which a competitor rides on a sled in a prone position and slides down a course of ice head-first."

4. Evaluation tests

We let Metis join two NTCIR 6 tests: Cross-Language Question Answering (CLQA) that handles factoid-type questions, and Question Answering (QAC) that deals with questions of the "why," "how," and "definition" types. The results from the respective tests are shown in Tables 4 and 5.

Table 4 Accuracy of Metis answers in CLQA

Right	Right + Unsupported	MRR
0.0900	0.1100	0.1100

Table 5 Accuracy of Metis answers in QAC

A Rank Accuracy	B Rank Accuracy	C Rank Accuracy	Total Accuracy
0.1030	0.0340	0.0340	0.1710

In many of the cases in which the system was unable to give the right answer, the reason was that it failed to obtain a knowledge sentence containing the right answer. With the Formal Run, Metis was often unable to obtain a relevant knowledge sentence since it conducted a search with the sentence segments, not with the morphemes. For example, there was a question, "What % of economic growth did China set up as its target for 1998?" To this question, the keywords, "the target in economic growth" was a single segment. Therefore, as discussed earlier in Section 3.3, we switched from segment searches to morpheme searches in order to use both "economic growth rate" and "target" as the keywords. Thus, we have made Metis capable of retrieving a sentence such as "They set up this year's target in the economic growth rate at 8% and also proposed a major organizational reform of reducing the national governmental ministries and agencies to three-fourths of what they are today." Also, in the case a keyword is

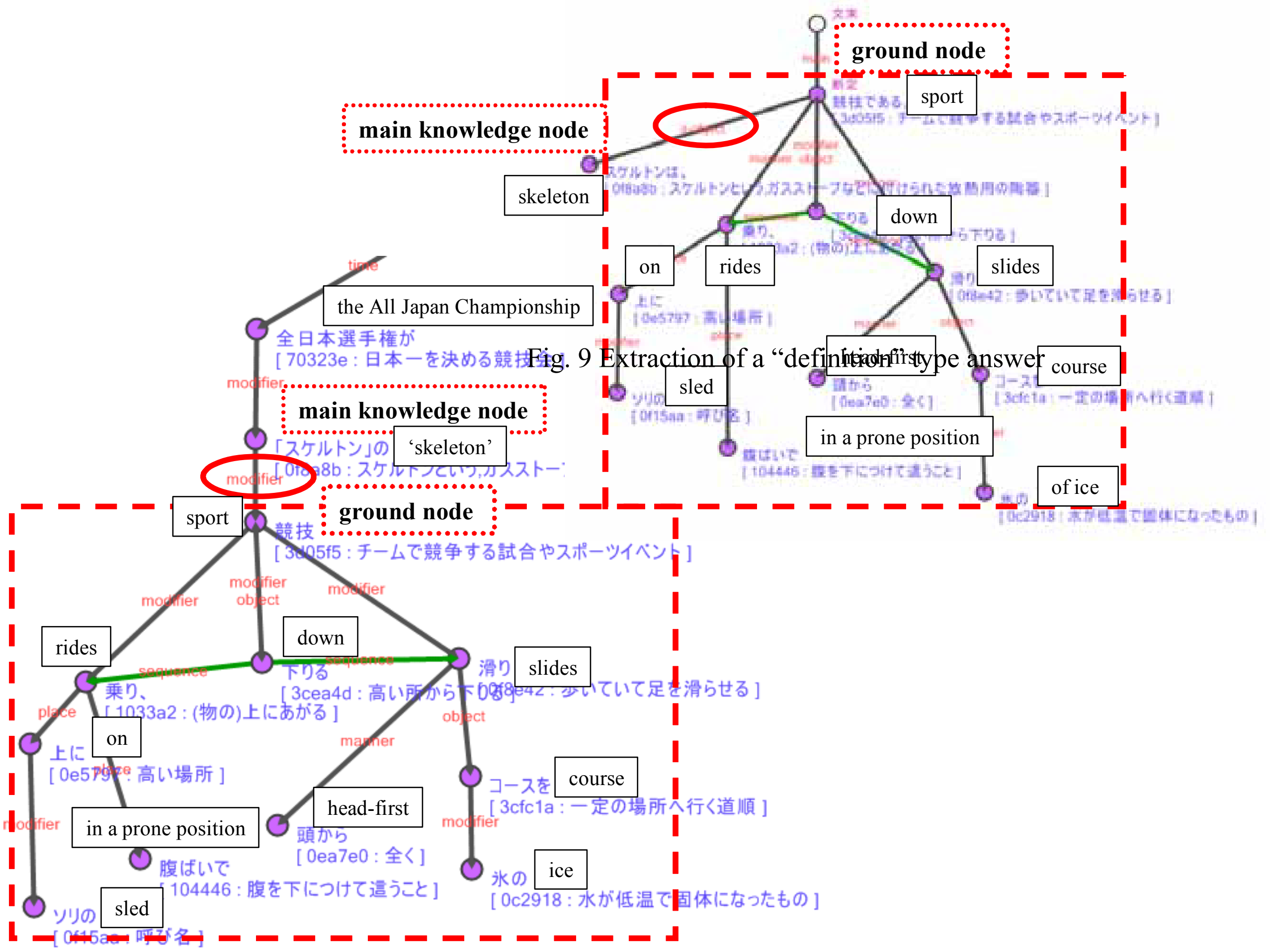

Fig. 9 Extraction of a "definition"-type answer

a Western calendar year, we have improved Metis so that it can search for both, for example, “1982” and “82,” and thus obtain more knowledge sentences.
Yet another reason for Metis to come up with a wrong answer was that the question and knowledge sentences had different syntactical structures that resulted in lower degrees of similarity between graphs. As described in Section 3.4, therefore, we added to Metis a function to paraphrase a question sentence so that the system can overcome expressional differences and conduct graph matching correctly. For example, to the question, “What country is Sampras, the male tennis player, from?”, the system now paraphrases it into “What nationality is Sampras, the male tennis player?” This way, the system is now able to match this question with the knowledge, “Sampras (US) won the championship at the male tennis.”
The system conducts a semantic analysis of a database of newspaper articles and uses the resulting semantic graphs. At the time of the Formal Run submission, due to time limitations, the system had not yet analyzed the half of the articles, resulting in some articles being unavailable as knowledge. After the improvements we added to it, however, the system semantically analyzed all of the articles contained in the database and now it is able to utilize all of the articles.

Tables 6 and 7 show the test results after the two improvements we added and the database update we made. The evaluation work was done manually, with humans evaluating the answers from the system.

Table 6 Accuracy of Metis answers in CLQA after the improvements

Right	Right + Unsupported	MRR
0.2750	0.2900	0.2206

Table 7 Accuracy of Metis answers in QAC after the improvements

A Rank Accuracy	B Rank Accuracy	C Rank Accuracy	Total Accuracy
0.1000	0.0349	0.0442	0.1791

Even after the improvements mentioned above, Metis was still unable to extract an answer in some cases, whose causes are described below:

Causes of inability to extract an answer:

- Appositions

Given the question, “Where is the capital of Kampuchea?”, in the case that the corresponding knowledge was written with an apposition, “Phnom Penh, the capital of Kampuchea,” Metis was unable to extract the answer since this appositional phrase was a single segment.

- Difference in katakana notations

When writing the name “Lewinsky” in Japan’s katakana, there can be several different notations to indicate the same name. Since Metis considers all such different notations to be different names, it sometimes retrieved some wrong sentences and made incorrect graph matching.

- Too much modification to a search keyword

In case the search keyword was the segment “Harvard University Kennedy School,” Metis ran a search with “Harvard University,” “Kennedy,” and “school,” since it conducts a search with morphemes. In a case such as this, a search with “Kennedy” and “school” alone, excluding “Harvard University,” can sometimes find some relevant knowledge. Similarly, there were some cases in which the system was unable to find a knowledge sentence due to too much modification of the keyword.

- Use of the date of issue of article

When answering to the question which specified the date like "What percentage was the unemployment rate of Japan in May, 1998?", because it is not inferable that May is May, 1998 when the article was written only with the knowledge "The unemployment rate of the man updated 4.3% and worst-ever though the ratio of complete unemployment in May remained in 4.1% in April which was worst-ever", the answer cannot be extracted.

Causes of incorrect answers extracted:

- Mistaken analysis of modification relation

Metis, as it matches graphs, pays good attention to modification relations and arc similarity, which is dependent on deep cases, in addition to degrees of node similarity. For this reason, if a dependency parsing analysis makes an error with a modification relation, Metis can place a segment irrelevant to the answer as the ground node to extract in the dependency tree. In some cases, the system extracted such a mistaken segment as the answer.

- Answers extracted out of a graph with a low degree of similarity

Since we put our emphasis on outputting answers to questions, we let Metis extract an answer out of a graph with a low degree of similarity as the result of graph matching. For this reason, sometimes the system produced incorrect answers. Though we can solve this problem by setting a higher threshold for graph matching, this solution can create another problem, i.e., the system becomes unable to provide an answer to some questions.

- Discrepancy between the requested answer and the corresponding deep cases

In cases involving extracting an answer to a question of the “why,” “how,” or “definition” type, Metis finds an answer based on the appropriate deep cases. Since we intended for the system to extract as many answers as appropriate, we let it choose more appropriate deep cases than are strictly necessary. In some cases, this has caused the system to provide an incorrect answer. One solution to this problem is to set up a score for a deep case in accordance with the situation, after carefully considering inter-word deep cases in knowledge sentences as well as the types of questions. Then, we can let the system extract cases with a score larger than a threshold value.

Processing to parentheses expression

Though it was inferable that the clause in parentheses modified the clause ahead of that when there was knowledge "Entitled <What were you able to buy by that money, bubble fantastic!> (Shogakukan)" for the question "What was the publisher of <What were you able to buy by that money?> written by Ryu Murakami?", the answer was not obtained from that knowledge because it cannot be judged that the clause in parentheses is a publisher and so lowered the degree of similarity between both sentences.

5. Conclusion

Metis is a system which extracts the solution of the question based on a semantic sentence matching, the algorithm is general in all the points of the retrieval, the matching, and the solution extraction, and a detailed tuning is not done. In the evaluation tests, the first answers held the largest share among the correct answers. This is evidence that Metis conducts precise matching of graphs. This precise matching, however, allows the system to be easily affected by the precision of semantic analysis. Thus, when matching long sentences, the system often made a mistake in dependency parsing, which in turn brought down the system’s precision. Because knowledge including the correct answer cannot often be retrieved, it is necessary to raise the retrieval success rate as a problem in the future. Moreover, to raise accuracy, a detailed tuning like the dependency analysis, apposition, parentheses, and the date etc. of the article is requested.

Acknowledgement

Part of this research was financed by a subsidy for “Research and development of a highly precise question-answering system that makes precise matching of a semantic base,” Basic Research C, for the scientific research of Japan’s Ministry of Education, Culture, Sports, Science and Technology. We hereby express our gratitude to the Ministry for their subsidy.

References

[1] Gakuto Kurata, Naoaki Okazaki, Mitsuru Ishizuka: Question answering system with graph structure based on dependency analysis, IPSJ-SIGNL, 2003-NL-158, pp. 69-74(2003).

[2] Masaki Murata, Masao Utiyama, Hitoshi Isahara: Question Answering System Using Similarity-Guided Reasoning, IPSJ-SIGNL, 2000-NL-135, pp161-188 (2000).

[3] Kazuaki Takehara, Kensuke Abe, Tomonari Yasuda, Dongli Han and Minoru Harada: A precise matching based on semantic analysis between the question sentence and knowledge sentence for question-answering, Proc. 66th IPSJ Annual Convention, 6U-03, pp. 2-173-174 (2004).

[4] Kazunori Sugimura , Tetsuya Yamamoto, Kentarou Kimura, Hayato Torii, Dongli Han, and Minoru Harada: Improvement of the precision of the semantic analysis system SAGE, and generation of Semantic Graph, Proc. 67th IPSJ Annual Convention, 1J-2, pp. 2-67-68(2005).

[5] Lucene.net http://incubator. apache. org/lucene. net/.

[6] Google Web APIs (beta) http://www. google. com/apis/.

[7] EDR http://www2. nict. go. jp/r/r312/EDR/

[8] Question Answering Challenge(QAC) http://www. nlp. is. ritsumei. ac. jp/qac/.

[9] Cross Language Question Answering (CLQA) http://clqa. jpn. org/.

[10] Juman, knp http://nlp. kuee. kyoto-u. ac. jp/nl-resource/top. html