The Interactive Navigation to the Stored Q&A data using Simple Questions

Kunio MATSUI
Fujitsu Laboratories Ltd.
matsui.kunio@jp.fujitsu.com

Hozumi TANAKA
Tokyo Institute of Technology
tanaka@cs.titech.ac.jp

Abstract

As an effort aimed at designing a system to save operator work at Call Centers, this research suggests directing users to Q&A data stored in the call center by using naturally expressed questions (simple questions) received from users as a trigger. And, based on actual data, this research effort closely examines difference in accuracy when selecting methods for key terms, and the difference in accuracy when case elements are omitted. Our efforts showed that the term that limits or extends the heads of the sentence greatly influences the search accuracy. Accordingly, this paper presents an interactive navigation method for supplementing a term in accordance with a rule and improves its effectiveness.

1 Introduction

Call Centers, which provide answers on product usage directly to consumers, are a rapidly growing part of customer service. However, due to the furious rate at which new products are being developed and introduced to the market, an ever-increasing knowledge base is required to effectively respond to user questions. This is making it extremely difficult for call center operators to give precise answers to complicated questions in a timely manner. Because of the stressful nature of this job, the stability of the workforce at Call Centers is decreasing and companies are facing an increasing number of problems in retaining a sufficient number of skilled operators, due to such costs as personnel expenses and employee training.

In an effort to design a system to assist Call Center operators, this report suggests an interactive navigation method that will channel callers to Q&A data stored in the Call Center database by using naturally expressed questions and natural sentences (hereinafter, simple questions). We believe that these simple questions can provide reliable indicators into the situation confronting the user and can be used as a trigger. In this paper we will also examine the difference in the precision of creating methods for simple questions, the difference in the precision of selecting methods for key terms, and the difference in the precision when case elements are omitted.

2 Related Research

While the ideal solution to the problem would be a computer that could answer user questions automatically, no such system is currently available and it is not clear as to when such a system will appear in the future.

As a research of dialogue, ELIZA (Weizenbaum, 1966) is very famous but it is not developed for the purpose of user solution. Also the interactive information retrieval system (Oddy, 1977) is designed to give alternatives to the user, but it is difficult to reflect the user intention.

As a research of Q&A system, a helpdesk system for software products (Kurohashi et al, 2000) is now working on research based on FAQ Finder

(Bruke, 1997), but it is still confined to a limited field.

At the same time, however, a method that uses previously stored question and answer data to provide responses to user questions is available. This method utilizes several approaches. One approach is a system where the operator mediates user inquiries and uses the system as a resource for finding the answer to the question. In this case, system integrity is not a requirement because of the human mediation. And some are effectively working in the sense of assisting operators even though they may be imperfect.

Alternatively, in some systems, the computer responds directly to the users. An example of such systems would be a system (Harabagiu, 2001) where a search question is replaced by creating a precise search query that may be used for QA task of TREC, and in another system(Robin, 1997) a planning question is replaced to clarify its goal by applying the world model and the user model in a limited world.

The problem with the former is that the question itself is not realistic and storing all the Q&A data costs a great deal. The problem with the latter is that creating the models costs too much for adopting to the rapidly changing real-world situations.

In yet another method, the stored Q&A data are utilized as FAQ to allow users to perform keyword search. However, a log survey of some internet search engine shows that 90% of users enter just one keyword for the first search step (Harada, 1997). This means it is hard to say that users are sufficiently utilizing this method as a search query even though there may be potential demands.

In light of the situation described above, the research effort described in this paper works at the method of solving user questions by generating a query that extracts effective constituent elements of simple questions and navigates the user to the stored Q&A, using "simple questions" that conform to the user's situation and that indicate the expression of natural questions.

3 Data and Evaluation Method

3.1 Q&A at Call Center

In this experiment, a group of records describing questions solved by operators at Call Center through telephone responses were divided into questions (<subject>) and answers (<solution>), and were used as an analysis target. An actual example is shown in Example 1.

The question and answer are not an exact dialogue record between the user and operator. It is considered that they actually began from an expression of a simple user situation or intention, and that the outcome derived from several dialogues between user and operator is summarized later. Data stored at Call Center are usually summarized in many cases.

```
<qaid>354</qaid>
<subject>
   I want to record and reproduce audio using the audio re-
corder with USB connection, Roland AUDIO CANVAS UA-
100. Please let me know connection method of putting the
audio outputs of PC together.
</subject>
<solution>
   For the audio recording and reproduction using Roland
AUDIO CANVAS UA-100, audio recording without getting
influence of the noise as AD/DA conversion can be sepa-
rated from the main body of PC (the sound card).
・Operating Procedure
   1. PC is connected with the terminal USB of UA-100.
(Recording and reproduction of WAV files)
   2. The terminal LINE OUT of the sound board on the PC
is connected with the terminal INPUT of UA-100. (Music CD
reproduction and reproduction of MIDI files that use sound
board or soft MIDI sound source)
   3. The terminal OUTPUT of UA-100 is connected to an
external speaker.
   4. In addition, audio recording can be done by connecting
the mike and the recording machine (MD and cassette, etc.)
with UA-100.
</solution>
```

Example 1 Sample Q&A at Call Center

To begin analysis of the stored data, we first conducted research on the constituent elements of questions. As the research viewpoint, 300 arbitrarily extracted questions were categorized into three types (Phenomenon, Denial, and Intention) as situation explanation, and nine question types (What, How, Method, Which, Why, Where, Request, Possibility, and Y/N), which were counted respectively. As the result below shows, the majority of questions consist of question sentences that can be expressed as situation explanation and question type. The majority consists of situation explanation and question type.

- Situation Explanation
   Phenomenon (154)
   Denial (92)

Intention (28)
- Question Type
   What (12), How (24), Method (20), Which (7),
   Why (17), Where (2), Request (47),
   Possibility (18), Y/N (21)

Such explanations and questions were not complete in the beginning. Instead, they were considered to have started with an expression of simple situation or question derived by the operators based on the dialogues with the users which were later summarized.

3.2 Setting of Goal

As previously described, when users call a Call Center, the process starts with simple questions in many cases consisting of situation explanations and conditions that are not complicated. Operators then ask the users for various details in order to lead them to the right answers. To perform the same type of navigation with a computer, it is first necessary to receive simple questions from users, and then to try to match those questions with stored questions by supplementing necessary conditions through interactive dialogues between the computer and the user, and then to lead the user to the right answer. (Figure 1)

```
User              Stored ===== Stored
Simple Questions   Questions   Answers

Possibility of dialogues for the matching process
```
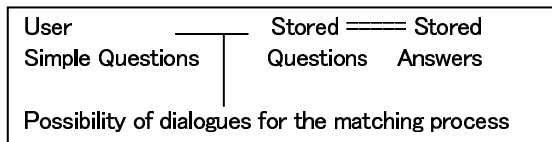
Figure 1 Image of the System Realization

In this experiment, we set a target for analysis on which points, and in what way, simple questions are matched with stored questions. Additionally, in an effort to find out what types of dialogues should be conducted for simple questions input by users, we looked to determine what kind of constituent elements in simple questions are effective for assisting matching.

For this, we decided the goal would be to first create simple question sentences that are not in the record, and then create a search query that best matches the stored questions.

3.3 Creation of Simple Questions

We created simple questions by reading through the stored questions and using various

user questions that can be analyzed as simple structure questions.

In order to avoid being too dependent upon creator's habits, here we tried, to the greatest degree possible, to create natural questions by indicating two methods and having plurality of persons involved in the work.

Two guidelines of creation are listed below:
- First guideline: "Natural and simple expressions where nature of the question is very easily recognizable from the sentence alone."(Ex. 2)

```
<simple_q>
<sq>I want to know the connection method to put audio
outputs together. </sq>
<sq>I want to record and reproduce audio.</sq>
<sq>I want to record and reproduce audio with ROLAND
AUDIO CANVAS UA-100.</sq>
<sq>I want to connect audio outputs all together.</sq>
</simple_q>
```

Example 2   Simple Question of Free Description

- Second Guideline: "Render the stored questions into several constituent elements making each element a simple question."(Ex. 3)

```
<simple_q>
<sq>I want to record and reproduce audio. </sq>
<sq>I want to use an audio recorder with USB connec-
tion.</sq>
<sq>Please let me know how to connect it with PC.</sq>
</simple_q>
```

Example 3   Simple Question of Constituent Elements

3.4 Retrieval Tool

As a retrieval tool, the preceding sentence search engine (Matsui et al, 2000), which we have developed, was used to retrieve the questions rendered from simple questions. Simple questions for the query were morphologically analyzed by JUMAN (Kurohashi and Nagao, 1999) to select key terms to be used for retrieval. Selection of key terms is shown below.

[Basis of selection of key terms used for retrieval]

After the morphological analysis of simple questions, key terms are set so that they can be selected by the following part of speech.
{Noun, Undefined Word, Adjective, Verb, Adverb}

Additionally, it can be selected from the level that expresses the head of case frame. For this, division and connection are defined as below, so that case frames are expressed. The charge parti-

cle and the case-marking particle, etc. at the level of the morphological analysis are recognized as divider, as the sentences are simple.

Divider : Divides the unity of each case frame.
 : が(ga), を(wo), で(de), に(ni), は(wa), も(mo), へ(e), "Punctuation"
Connector : Expands the range of each case frame.
 :の(no), "Noun Connection"
Nouns : Noun, Undefined Word

Using the definition of these connectors and dividers, the levels of the key terms were defined as follows. The levels were defined for the purpose of determining whether interactive key terms supplementation was necessary.

Level 1 : Only the term immediately preceding a divider
Level 2 : All terms that are connected by a connector preceding a divider.
Level 3 : All terms

By selecting a speech part and level, key term assignment becomes possible as described below.

Example 4 Output of morphological analyzer and the dividers

[Sample sentence]
留守番電話で番号を通知させる方法がわからない.
(rusubandenwadebangouwotsuuchisaseruhouhougawa-karanai.)
[Output of morphological analyzer]
rusuban / denwa / de / bangou / wo / tsuuchi / sase / ru / houhou / ga / waka / ra / nai /.
(I don't know how to have the number displayed on the display screen of the answering machine.)

The dividers were underlined in the above-mentioned of morphological analyzer output. In this example, the following terms can be assigned at each level.

Level 1 : 電話(denwa), 番号(bangou), 方法(houhou), ない(nai)
Level 2 : Level 1 + 留守番(rusuban)
Level 3 : Level 2 + 通知(tsuuchi), させ(sase), わか(waka)

Questions that have accumulated were pre-registered as features of n-gram when the indexes were made. The general weighting method of TF/IDF(Term Frequency and Inverted Document Frequency) was used to put the retrieval output in order of relevance.

3.5 Evaluation metrics

Since its aim is to navigate to the accumulated questions, measuring the recall ratio and the relevance ratio used in a general retrieval was not considered a valid alternative. Hence, evaluation was conducted based on the order in which the ranking was done with the retrieval tool.

Relative evaluation was done according to the number of first places.

4 Experiments

The experiments were conducted by the following methods.
[Methods]
- Ranking search was conducted on simple questions as queries and accumulated questions as targeted information.
- Ranking search was conducted on question and answer pairs, with a combination of plural, simple questions designated as queries.
- Number of queries: 300 Queries, made as simple questions were randomly extracted from 38,000 accumulated questions.
- Number of accumulated questions: 38,000.
[Conditions]
- Difference by basis of selection of key terms
- Difference by excluding case elements
- Supplement of term to simple question
- Retrieval method: The morpheme was analyzed with JUMAN. With a noun, a verb, an adjective, an undefined word, the prefix, the suffix, and the adverb assigned as key terms, retrieval was conducted as follows
.
L1-one: a single, simple question and the key terms of level 1.
L1-all: plural simple questions and the key terms of level 1.
L2-one: a single, simple question and the key terms of level 2.
L2-all: plural simple questions and the key terms of level 2.
L3-one: a single, simple question and the key terms of level 3.
L3-all: plural simple questions and the key terms of level 3.

Next, the retrieval type, without the "case element" delimited by a divider, was made for the level 3 key terms, and retrieval was conducted as follows.
one-ga: a key term excluding "ga" case with a single, simple question

all-ga: a key term excluding "ga" case with plural, simple questions

one-wo: a key term excluding "wo" case with a single, simple question

all-wo: a key term excluding "wo" case with plural, simple questions

one-de: a key term excluding "de" case with a single, simple question

all-de: a key term excluding "de" case with plural, simple questions

one-wamo: a key term excluding "wa" case and "mo" case with a single, simple question

all-wamo: a key term excluding "wa" case and "mo" case with plural, simple questions

one-nie: a key term excluding "ni" case and "e" case with a single, simple question

all-nie: a key term excluding "ni" case and "e" case with plural, simple questions

5 Arguments on the composition of simple questions

5.1 Difference in method of making simple questions

Simple questions were made using two guidelines. Along with its comparison, the difference between a retrieval type made only from a single, simple question and a retrieval type made by using the logical sum of plural, simple questions are shown below. Table 1 shows the retrieval output using simple questions made by the first guideline (hereafter, sqf), and Table 2 shows the retrieval output using simple questions made by the second guideline (hereafter, sqc).

Table 1   Retrieval output using the 1st guideline (sqf)

Ranking	L3-one	L3-one	L3-all	L3-all
1st	414	53.6%	232	77.3%
2nd-10th	184	23.8%	44	14.7%
below 11th	175	22.6%	24	8.0%
	773	100.0%	300	100.0%

Table 2   Retrieval output using the 2nd guideline (sqc)

Ranking	L3-one	L3-one	L3-all	L3-all
1st	248	49.7%	223	74.3%
2nd-10th	115	23.0%	51	17.0%
below 11th	136	27.3%	26	8.7%
	499	100.0%	300	100.0%

From the results of Table 1 and Table 2, it is clear that retrieval using combined plural, simple questions yields a higher ratio of first place in the retrieval order than retrieval by a single, simple question. Since both outputs were 24-25% higher, it is understood that, from the viewpoint of retrieval, a simple question does not increase noise, but rather, proves itself to be an effective key term. In other words, a single, simple question alone cannot easily navigate users to the original question, it requires supplementary information.

In addition, the retrieval outputs by different guideline of constructing simple questions showed that the first guideline gave higher ratios of first place by about 3-4% in both single and plural questions. This is because there seemed to be a tendency for … to satisfy the accumulated original questions even though simple questions were freely described in the first guideline. Example 5 are samples of simple questions that led to the retrieval output in which the simple question made by the second guideline was significantly worse than the simple question made by the first guideline

In <qaid>769 of example 5, "sold at distribution outlet" is the important key term, and in <qaid>1510, "built-in fax modem" is the important key term. As described later in the method of selecting the key term in which similar results are seen, an important key term for retrieval is included in words and phrases that modify nouns.

<qaid>769</qaid>
<subject>
FM-TOWNS FRESH) WINDOWS95 bought in a distribution outlet cannot be installed. Why?
</subject>
<sqf>> WINDOWS95 bought in a distribution outlet cannot be installed </sqf>→1st rank
<sqc>> WINDOWS95 on the market cannot be installed .</sqc>→271st rank

<qaid>1510</qaid>
<subject> The following errors occur when dialed to the Internet with the modem by using dial-up networking at the connection destination. Show what to do. Also, show how to turn off the power supply of the built-in fax modem in software.
</subject>
<sqf>> I want to know software-wise the method of turning off the power supply of the built-in fax modem.</sqf> →1st rank
<sqc> Cannot dial-up-connect with the modem. </sqc> →29th rank

Example 5   Difference between sqf and sqc

5.2 Difference by basis of selection of key terms

Table 3 shows the retrieval output with change in the selection criterion of the key terms for the above-mentioned sqf. Likewise, the retrieval output with a change in the selection criterion of the key terms for sqc is shown in Table 4.

What can be said about both sqf and sqc is, Level 1 that selects only a term just before a divider has lower ratio of first place by more than 30 points compared with Level 2 (terms connected by connectors) and level 3 (all terms). In the Japanese language, a term preceding a particle usually becomes a head, which indicates that a head alone is not sufficient as a key term for retrieval.

Table 3 Difference by basis of selection of key terms (sqf)

Ranking	L1-one	L2-one	L3-one	L1-all	L2-all	L3-all
1st	15.7%	47.2%	53.6%	34.7%	69.0%	77.3%
2nd-10th	21.1%	24.6%	23.8%	25.3%	19.3%	14.7%
below 11th	63.3%	28.2%	22.6%	40.0%	11.7%	8.0%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 4 Difference by basis of selection of key terms (sqc)

Ranking	L1-one	L2-one	L3-one	L1-all	L2-all	L3-all
1st	15.2%	41.1%	49.7%	32.7%	64.0%	74.3%
2nd-10th	22.4%	24.2%	23.0%	28.7%	21.3%	17.0%
below 11th	62.3%	34.7%	27.3%	38.7%	14.7%	8.7%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

5.3 Difference by excluding case elements

Table 5 shows the effect of the removal of the case element in a single, simple question for sqf, and Table 6, the effect of the removal of the case element in plural, simple questions for sqf. Likewise, Table 7 shows the effect of the removal of the case element in a single, simple question for sqc, and Table 8, the effect of the removal of the case element in plural, simple questions for sqc. The effect is most remarkable when "ga" case element or "wo" case element is removed for both single and plural questions: both ratios of first places have decreased by more than 15 points. In addition, the ratio of first places has decreased by 6-7 points in "ni, e" case element. However, for sqc, there is little change in "de" case element as the situation differs between sqf and sqc. This is due to the difference in writing a simple question.

Table 5 Effect of the removal of the case element in a single, simple question (sqf)

Ranking	L3-one	one-ga	one-wo	one-de	one-hie	one-wam
1st	53.6%	37.8%	35.8%	47.9%	47.3%	52.3%
2nd-10th	23.8%	19.5%	20.2%	25.1%	23.3%	23.9%
below 11th	22.6%	42.7%	44.0%	27.0%	29.4%	23.8%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 6 Effect of the removal of the case element in plural, simple questions (sqf)

Ranking	L3-all	all-ga	all-wo	all-de	all-nie	all-wamo
1st	77.3%	63.7%	61.7%	71.3%	70.0%	76.0%
2nd-10th	14.7%	17.7%	20.0%	18.0%	17.0%	16.0%
below 11th	8.0%	18.7%	18.3%	10.7%	13.0%	8.0%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 7 Effect of the removal of the case element in a single, simple question (sqc)

Ranking	L3-one	one-ga	one-wo	one-de	one-hie	one-wam
1st	49.7%	34.7%	36.5%	46.3%	41.5%	45.7%
2nd-10th	23.0%	19.6%	20.4%	23.0%	22.2%	23.2%
below 11th	27.3%	45.7%	43.1%	30.7%	36.3%	31.1%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 8 Effect of the removal of the case element in plural, simple questions (sqc)

Ranking	L3-all	all-ga	all-wo	all-de	all-nie	all-wamo
1st	74.3%	55.3%	58.0%	70.3%	63.3%	69.3%
2nd-10th	17.0%	19.7%	20.7%	18.3%	21.0%	18.3%
below 11th	8.7%	25.0%	21.3%	11.3%	15.7%	12.3%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

6 Discussion concerning interactive navigating

6.1 Strategy of supplement of interactive term

It remains in about 50 percent in a single, simple question though the nearly 80 percent of questions can be navigated by combining two or more single, simple questions. It is important to determine whether it is possible to navigate to supplement a term very appropriate to surroundings of a head as being from the discussion about Chapter 5 in order to improve the search precision.

It is necessary to decide which term in the simple questions attention should be focused on, and whether to limit the term or to provide it in detail. What query the system makes becomes the strategy of the term supplement.

Consequently, to examine whether the kind of the noun relates as a clue that decides whether to limit the term, or to make it in detail, about 38,000 accumulated questions were analyzed. We then counted the number of the questions in

which the modification by "no" appears five times or more.

Tables 9-11 show the example of typical phrases, the total frequency, and the ratio. Table 9 shows the relation between "sahen" noun (Japanese verbal noun) and "no". Table 10 shows the relation between the common noun and "no". Table 11 shows the relation between the proper noun and "no".

Table 9: Relation between "sahen" noun and "no"

Term	Fregency	modify with "no"	modified by "no"
set up	722	32	690
display	166	9	157
install	158	16	142
update	144	4	140
connect	119	18	101
Total	5107	862	4245
Ratio		16.9%	83.1%

Table 10: Relation between common noun and "no"

Term	Fregency	modify with "no"	modified by "no"
file	262	132	130
data	182	75	107
hard disk	121	85	36
PC	107	77	30
program	96	47	49
Total	14181	6362	7819
Ratio		44.9%	55.1%

Table 11: Relation between proper noun and "no"

Term	Fregency	modify with "no"	modified by "no"
Windows95	180	149	20
Windows98	85	76	9
Windows	65	61	4
Word	37	35	2
Fujitsu	36	35	1
Total	1567	1405	162
Ratio		88.9%	11.1%

6.2 Supplement of term to simple questions

The following four rules, that decide which term in the simple questions attention is to be focused on, and how to limit it or make it in detail, can be derived in consideration of the above analysis.

(1) Focus attention to the term that becomes a head of the case frame in the order of the case frame that influences the search precision. ("ga" > "wo" > "ni/e" > "wa/mo" > "de" as the result of Table 7 and 8)

(2) Focus attention to the term of "A" when the term that becomes a head contains the connector and is shape of "B of A" (because B has already been limited).

(3) Supplement the term that limits the "sahen" noun if the term that is focused attention is a "sahen" noun (as the result of Table 9).

(4) Supplement the term that defines the proper noun in detail if the term under intense consideration is a proper noun (as the result of Table 11).

By using these rules, the term under intense consideration and the response of system according to the following procedures, and determine if the term should be supplemented.

[Procedure of supplement of term]

(a) When the case frame divided by divider except "Punctuation" is plural, rule (1) is applied.

(b) Table 12 shows that the response of the system changes depending on the presence of the existence of the connector "no" and the kind of the noun before and behind. The upper row shows the number of the applied rule. The lower shows the response of the system. The sign "|" shows that either possibility exists. Moreover, as a heuristics, "Donna (Which)" is used for the limitation of the "sahen" noun, and "Nanno (What)" is used for the limitation of the common noun.

Table 12: Term in "B of A" that is focused attention

A\B	–	Sahen	Common	Proper
Sahen Noun	(3)	(2)(3)	(2)(3)	(2)(3)(4)
	「"Donna" A」	「"Donna" A」	「"Donna" A」	「"Donna" A」| 「B"no Nani"」
Common Noun	×	(2)(3)	(2)	(2)(4)
	「"Nani no" A」|「A"no Nani"」	「"Nani no" A」|「"Donna" B」	「"Nani no" A」	「"Naini no" A」|「B"no Nani"」
Proper Noun	(4)	(2)(3)(4)	(2)(4)	(2)(4)
	「A"no Nani"」	「A"no Nani"」|「"Donna" B」	「A"no Nani"」|「"Nani no" A」	「B"no Nani"」

```
<qiad>351</qaid>
<sqc> Should I set CRT? </sqc>.
<sqc-new> Should I set CRT for FMV-5133T3? <sqc-new>
<qa1d>384</qaid>
<sqc> The error occurs in "FUDEMAME." </sqc>
<src-new> The error occurs in the address list of "FUDEMAME." <sqc-new>
```

Example 6: Examples of new simple questions

At this time, the example of supplementing the term is shown in example 6. (parts enclosed with <sqc-new> are new simple questions).

The simple questions, which were low in rank when searched by the simple questions of sqc shown in Table 2, were reviewed along this flow.

Table 13: Search result by this new simple question (sqc-new)

Ranking\Condition	by original simple questions				by new simple questions			
	L3–one		L3–all		L3–one		L3–all	
	Num.	Rate	Num.	Rate	Num.	Rate	Num.	Rate
1st	248	49.7%	223	74.3%	346	69.3%	269	89.7%
2nd–10th	115	23.0%	51	17.0%	109	21.8%	24	8.0%
below 11th	136	27.3%	26	8.7%	44	8.8%	7	2.3%
	499	100%	300	100%	499	100%	300	100%

Then, one term was added to the original simple question, and we experimented to determine how much the search precision had improved. Table 13 shows the comparison result. In Table 13, the ratio that wins the first place rose to about 20% in case of a single question and to about 15% in case of plural questions respectively. As a result, the precision improvement of confirmed.

7 Future Work

The future works are shown below.
(1) Absorption of difference of phrasing
One of the causes of inability to navigate is a mismatch of terms by the difference of notation or expression. To correct this, the search by ontology is needed.
(2) Dealing with compound clause and embedded clause
It is necessary to deal with compound clauses, and embedded clauses, even though all the simple questions used in this research were simple sentences.
(3) Verification by mass data and real data
This time only the amount of about 10% of those simple questions made were targeted due to the limit of processing time. It is necessary to put out the result of no bias by all simple questions. Moreover, it is necessary to obtain the first question that a user asks on an actual scene, and to verify it by real data though simple questions were made this time.

8 Conclusion

In this experiment, a method was considered, in which the expression of a simple doubt was replaced with a simple question, and then previously recorded question and answer was retrieved. We now understood that the terms that limited a head or made it in detail were required when navigating to the solution that the user is seeking. And it was proposed how to draw out those terms interactively, and its effectiveness was proven.

References

Weizenbaum, J. 1966, ELIZA –A computer program for the study of natural language communication between man and machine. Communications of the ACM, 9(1) 36-45

R.N.Oddy, 1977. Information Retrieval through Man-Machine Dialogue, The Journal of Documentation, Vol.33, Number 1

Sadao Kurohashi and Wataru Higasa. 2000. Dialogue Helpsystem based on Flexible Matching of User Query with Natural Language Knowledge Base, In Proceedings of 1st ACL SIGdial Workshop on Discourse and Dialogue, pp.141-149

Robin D. Bruke. 1977, Question Answering from Frequently-Asked Question Files: Experiences with the FAQ Finder System, Technical Report TR-97-05

Sanda M. Harabagiu. 2001, Just-In-Time Question Answering, NLPRS-2001, pp. 27-34

Allen J.F. 1983. Recognizing Intentions from Natural Language Utterances in Computational Models of Discourse, MIT Press, pp.107-166

Harada, Shimizu. 1997, Utilization of User Log on the WWW Search Engine, IPSJ, 97-DPS-81, pp.61-66

Kunio Matsui, Isao Namba, Nobuyuki Igata. 2000, Full-text Search Engine, The Technology of Information and Science, Vol.50 No.1, (2000-1)

Sadao Kurohashi, Makoto Nagao. 1999. JUMAN version 3.61(Manual), Kyoto University