Structuring Utterance Records of Requirements Elicitation Meetings Based on Speech Act Theory

Motoshi Saeki† Kinji Matsumura† Jun'ichi Shimoda† Haruhiko Kaiya‡

†Dept. of Computer Science, Tokyo Institute of Technology
‡School of Information Science,
Japan Advanced Institute of Science and Technology, Hokuriku

Abstract

This paper discusses a technique to structure utterance records of the meetings for requirements elicitation based on speech act theory. To elicit requirements to the system to be developed, the customers and the developers often have a series of face-to-face meetings. Utterances in verbal conversation made during the meetings include not only declarative information but also speech acts such as "commitment", "supporting a solution", "explaining a rationale" and so on. Extracting this information leads to clarifying customers' intent, decision rationale, and what problems have not been solved yet and remain unsolved in requirements elicitation phase. The point is how we extract speech acts from utterances and hold structurally them so as to use them afterwards. Our emphasis is on specific words (called keywords) included in utterances. These keywords can suggest what topics are discussed in the utterances and which speech acts the utterances have. Our technique has been assessed by experimental case studies and we have their good applicability to actual meetings for requirements elicitation.

1 Introduction

Software development is cooperative work performed by various kinds of persons, e.g. customers, users, project managers, requirements analysts, designers, programmers and so on. Especially, a requirements analysis phase, which is the earliest phase in a development process, has significant effect on developing high-quality software efficiently. We have several methods for requirements elicitation such as scenario analysis[12] and for specifying requirements such as structured analysis[4] and object-oriented analysis[15, 13], and tried to put them into practice[10]. The essential difficulty to complete requirements elicitation results from understanding and communication gaps between customers and analysts. A requirements elicitation phase often has two types of human activities ; the activities in *face-to-face* meetings which customers, users, analysts (interviewers) and so on may participate in (called meeting activities), and the activities of describing specification documents outside the meetings. These activities are usually repeated alternately and the specification is incrementally constructed. Verbal conversation in these meeting activities contains very useful information such as decision rationale which is not written explicitly in specification documents. More concretely, utterances in conversation contain not only descriptive information but also speech acts[14] such as "commitment", "supporting a solution", "explaining a rationale" and so on. This information is very important and necessary to elicit exact requirements from meeting activities. The point is how we extract speech acts from utterances and hold structurally them so as to use them afterwards.

Participants in the meetings sometimes misunderstand which topics have been decided as a conclusion. That is to say, they have identification gap to a conclusion. They cannot remember the discussion process how to come to a conclusion and it leads to increasing misunderstanding conclusions. It is very important to record the discussion processes and to access them during the meetings and when the participants describe the documents after the meetings. In particular, to avoid misunderstanding conclusions, we should hold structurally utterance records associated with the speech act information. From maintenance view, the traceability of requirements is one of crucial factors[5, 7]. Holding utterance records structurally and accessing them during a development process allow us to trace backward the life cycle of the requirements, i.e. to know why and how the requirements were created.

This paper discusses how to extract significant information such as speech acts from the utterances and how to structure them based on the extracted information, so that we can get useful information during meetings and on composing requirements specification. Our emphasis is on specific words (called *keywords*) included in utterances. By using the keywords extracted from an utterance, we identify

1. which parts of the system to be developed the utterance refers to and
2. which speech acts are included in the utterance.

The former information can be used to partition a sequence of utterances into segments which refer to a common topic. By using this information, we can access the utterance records with a topic.

The organization of the paper is as follows. In the next section, we will introduce our basic idea, i.e. what information can be extracted from an utterance sequence by using keywords. In addition, related work

will be referred to. Section 3 will present what structure and how we can construct from an utterance sequence by using keywords. Our technique consists of two steps — one is for partitioning the records into segments based on which topics are discussed in them, and the other is for identifying speech acts of the utterances and referential relationships among them. Domain-specific words denoting topics are used for segmentation in the first step, while keywords denoting speech acts are used in the second step. An example in the section helps readers to understand our technique. We will also show how useful attaching speech-act tags to recorded utterances are, by the example. In section 4, we apply our technique to experimental meetings in order to assess it. The result will be also discussed. Section 5 will clarify the directions of future work.

2 Basic Idea

In this section, we show the basic idea, i.e. what information is needed to structure utterance records and how we can extract it.

Customers or users and developers usually have requirements elicitation meetings at several times. The participants in a requirements elicitation meeting are not so many, e.g. four or five, less than ten. So a few of them seating side by side don't discuss locally the specific issues different from what all of them are discussing. All of them discuss and focus on an issue at any time. In the first meeting, the customers explain their initial requirements. After the meeting, the developers construct their solution to the requirements and present it to the customers in the next meetings. After that, the customers and the developers may have several meetings until they come to agreement. Meetings may have several phases — the phase where the customers explain their requirements to the developers, the phase where the developers (the analysts) present their solutions to the customers, the phase where the participants discuss the solutions to come to agreement, the phase for negotiating development schedule, and so on. Even a meeting can be divided into these phases. The characteristics of the participants' utterances depend on the phase that they are in. Thus, for each phase, we should design an individual structure to hold the utterances. In the case of a customer's explanation phase, as Colin Potts et al. pointed out in [12], the customer presents his or her requirements and the developers ask questions to clarify their unclear points. In such a phase, we will provide question-and-answer oriented structure to hold the utterance records. We should design the different structures for the other phases.

Each participant has a role such as "customer", "analyst" and "user" etc. Their behavior including giving an utterance depends on their roles and can be modelled as state transition machine. Each participant has internal states and his or her internal state changes by receiving an utterance from the others. A speech act can be considered as an abstraction of the utterances that cause the same state-transition pattern. For example, assume that a participant gives an utterance whose speech act is "proposal". After the hearers receive it, they change their internal states to the state where they should understand his or her proposal and express their positions clearly. Speech acts and state transitions depend on the phases where the participants are. Identifying speech acts of utterances is very helpful to extract their speakers' intent.

Next, we should consider how to identify a speech act of an utterance. An utterances may include specific words, i.e. keywords that can play an important role on identifying its speech act. Suppose that a participant gives the utterance "Or shall we ... ?". From the keyword "or" and the key phrase "shall we", we can identify that the speaker proposes an alternative [1]. It means that we can identify a speech act from a part of the utterance, i.e. keywords and key phrases included in an utterance, instead of the whole of the utterance.

We can have typical patterns of the sequences of speech acts occurring in conversation. For example, it is usual that immediately I send you the speech act "answering" after you give the speech act "questioning" to me. We can get a typical pattern of speech acts "questioning – answering". A set of pre-extracted typical patterns helps us to identify a speech act from the adjacent utterances.

In addition to speech acts, information about what an utterance refers to is very important. Even if we can identify that an utterance have the speech act "proposing", it is not so useful until we identify what the proposal is concerned with. We can also get this information by focusing on specific words appearing in the utterance. In particular, if we find domain-specific words in the utterance, we can guess that it refers to the corresponding constituents of the system to be developed. Suppose that a participant say "Or shall we use a image scanner?". We pick up the specific word "image scanner" and understand that this utterance is concerned with a "image scanner" part of the system.

We list up what information we extract from utterance records to structure them in the following;

- what speech act does an utterance has?
- which parts of the system does it refers to?

Some researchers have studied the application of speech act theory to modelling information systems. In the techniques[1, 11, 9], communication in information systems is classified by using speech act theory and then their behavior is modelled. Coordinator[16] is a structured E-mail system based on speech act theory. It has pre-defined transition patterns of speech acts and prescribes to mail senders and receivers what action they should perform. That is to say, they should follow the patterns. It causes users to the restriction of their actions and reduces the varieties of their actions. In requirements elicitation, which is a creative task, restricting the participants' actions is undesirable. In our approach, we never force the participants to perform restricted actions following speech act transition patterns. We use the transition patterns to extract speech acts from the utterances that

[1] The sentence "Or shall we ... ?" is a direct translation of a Japanese sentence to English. Thus we do not assert that the occurrences of both English word "or" and phase "shall we" always denote "proposing an alternative". In the case of Japanese, we can assert that the corresponding Japanese words express "proposing an alternative".

the participants give without any restrictions. Inquiry Cycle Model[12] is similar to our technique in using speech act theory. It captures requirements elicitation as inquiry-answer cycle and emphasizes classification of "questioning". To deal with the other speech acts and with other phases, the enhancement would be needed. Goguen and Linde discussed the applicability of discourse analysis to requirements elicitation[6]. Performing discourse analysis completely including conversation analysis needs much human efforts. Our technique can be considered as a variation of conversation analysis. However performing it has much lower cost because of focusing just on specific words in utterances.

3 Structuring Utterance Records

3.1 Overview of Our Technique

First of all, we should mention how to record conversation, i.e. sequences of utterances made in requirements elicitation meetings. Each participant has a microphone set and it is connected to his workstation to record his utterances electronically in the workstation. We can identify who give an utterance because a microphone is allocated to a participant. When the participant stops talking in a constant time period, e.g. 2 seconds, it is identified as a punctuation point of utterances.

Current techniques on speech recognition and natural-language understanding do not enable computers to recognize human conversation automatically yet. In this paper, first of all, we write down transcriptions with text from recorded conversation. We call them utterance text and we extract significant information from them. We will discuss the possibility of automatic processing by using word spotting technique in section 5.

Figure 1 shows the structure of the information that we extract from utterance text in Entity Relationship Diagram form. The solid round boxes and the solid lines in the figure denote the parts that the technique mentioned in this paper supports.

Assume that we have a set of keywords (incl. domain-specific words) and typical patterns of speech acts. Our technique is shown in the following:

1. identifying topics which an utterance sequence refers to by using domain-specific keywords and partitioning the utterance sequence into segments based on the referred topics. Each segment may consist of an utterance or of several utterances where a specific topic is discussed (Figure 2).

2. identifying speech acts of the utterances in each segment by using the keyword set and typical speech act patterns.

3. identifying referential relationships among utterances ("related_to" in Figure 1) by the typical speech act patterns.

3.2 Example

To explain our technique, we use an example of a 39 minutes' meeting where an analyst presents his solution to two users. In the example, the system to be developed is a hyper media system. First, we identify topics which the participants discuss in utterances

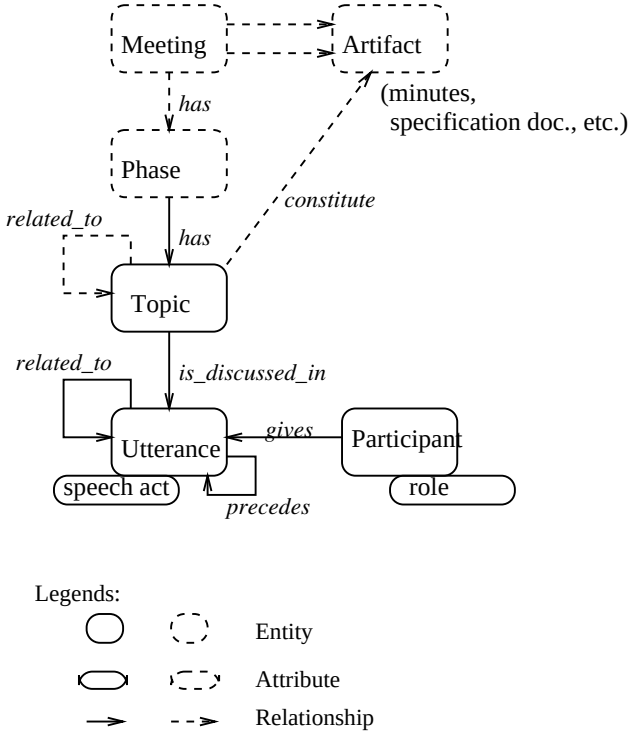

Figure 1: Structure of Utterance Records

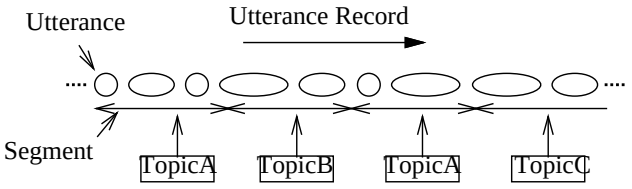

Figure 2: Partitioning Utterance records to Segments

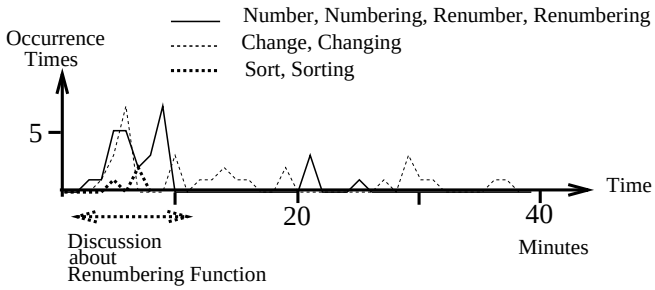

Figure 3: Occurrences of Domain-Specific Words in Utterance Records

and partition the utterances into segments. We focus on the function of renumbering node identifiers in the hyper media system as an example. We pick up the words "number", "renumber", "change" and "sort" and their inflections as domain-specific keywords denoting the renumbering function. Figure 3 shows the occurrences of the domain-specific words and its vertical axis expresses how many times the keywords occurred in one minute interval. For example, the words "number", "renumber" and their inflections occurred at five times in the interval between four and five minutes after the meeting started. Furthermore the keywords we picked up often occur in the interval from four to ten minutes. We could find that the participants discussed the topics concerned with "number", "renumber", "change" and "sort". We had this interval as a segment where they discussed the function of renumbering identifiers. As this example suggests, we can identify what topics are discussed in which parts of the utterance records by exploring the occurrences of domain-specific words.

Figure 4 shows the utterance text of this segment. Note that we have translated Japanese utterances directly into the English text shown in the figure, so it may not be natural in English native conversation. However, the direct translation (word-to-word translation) is needed to explain the steps of our technique because the emphasis is on specific words included in utterance text.

Next, we focus on the keywords which help us identify a speech act of an utterance. The speech acts we use in the example are "proposing", "explaining details (incl. rationale)", "questioning", "answering yes", "answering no", "answering others (neither yes nor no)", "agreement" and "disagreement" because the meeting is in the phase where an analyst presents his solution to the requirements and the users agree or disagree to them. These speech acts allow us to identify which topics remain unsolved problems without agreement.

We can easily identify that the utterance #1 explains rationale because it contains the specific keyword "because". In the example of the utterance #2, we can identify that it has the speech act "Questioning" because of the occurrence of the key phrase "let me ask a question". Similarly the next utterance, i.e.

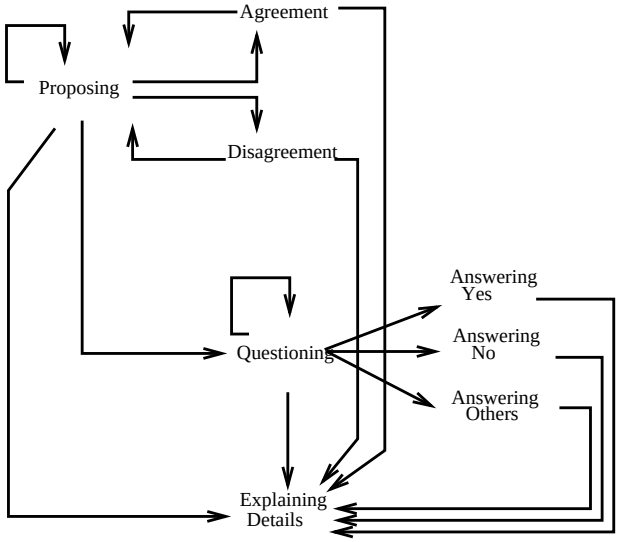

Figure 5: Typical Pattern of Speech Act Transition

#3 can be recognized to be "Answering Yes" to the utterance #2. We should mention the difference between English and Japanese again. In Japanese, unsimilar to English, questioning sentences can be easily identified because they contain specific suffix words expressing questions. The suffix words cannot be directly or explicitly translated into English, so they do not appear in the Figure 4. We can identify which utterances express question very easily in Japanese. In the case of the utterance #5, it is not so easy to decide that the utterance is a question only by focusing specific words if we scan it word by word as an English sentence. However it is easy in the case of an original Japanese sentence because the suffix word is included in it. The original utterance is "sosuruto (then), sorting (sorting) mitaina koto (function) wo suru wake (have) desu ka ?" and "ka" is the specific word denoting a question.

Figure 5 shows a typical pattern of speech act transition in the analysts' presentation phase. It is used for identifying speech acts of the utterances that could not be decided by the keywords and for identifying referential relationships between utterances. Assume that a participant gives "proposing". According to the figure, another participant or he typically gives "(another) proposing", "agreement", "disagreement", "questioning", or "explaining details" to the "proposing". Concentrate on the utterances #3 and #4 whose speech acts are "answering yes" and "proposing". We did not have the typical transition from "answering yes" to "proposing" as shown in Figure 5. It suggests to us that the utterance #4 might not have referential relationship to the #3. Finally we decide which utterance it has relationship to if any. The typical transition patterns help us to decide the relationships. Of course, however the final decision should be made by human workers.

Finally we can construct the structure shown in

No.	Speaker	Speech Act	Utterance Text
1	Analyst	Explaining Details (Rationale)	I would like to be able to improve the display layout because your idea is not so good for displaying final contents afterwards.
2	User	Questioning	Let me ask a question. Can we change the topic node numbers?
3	Analyst	Answering Yes	Yes, you can change the numbers. When you delete a topic node with a number, a gap of numbering is created. At that time, the system can renumber the nodes to pack the gap. I will design the system having this function.
4	Analyst	Proposing	It is also good that users can change the numbering.
5	User	Questioning	Then, can the system have the function of sorting?.
6	Analyst	Answering Yes	Yes, I think so. I don't know that it's possible or not yet. I will try.

Figure 4: Example of Utterances

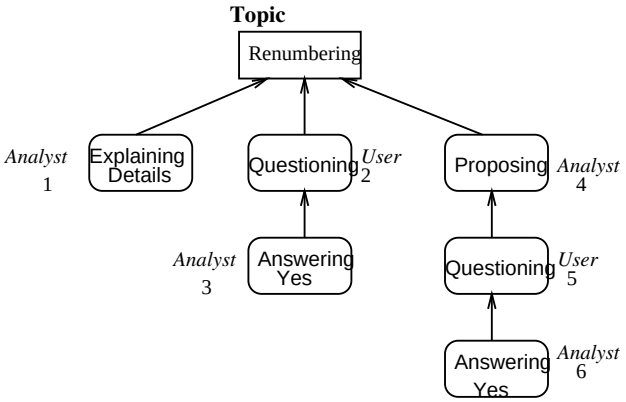

Figure 6: Extracted Structure

Figure 6 from this utterance sequence.

3.3 Using Recorded Utterances

In this section, we will illustrate how to use the structurally recorded utterances to clarify the advantages of our technique in a practical setting.

Utterances in conversation contain not only descriptive information but also speech acts such as "commitment", "supporting a solution", "explaining a rationale" and so on. This information is very important and necessary to elicit exact requirements from meeting activities. We will show how the information is important and necessary by the example.

Suppose that Figures 7 (a) and (b) are utterance texts which were given in different situations after the utterance in Figure4. In both of them, the participants are discussing the function of the renumbering the nodes. In Figure7 (a), as they could not remember the reason of adopting the renumbering function they gave up including it. As a result, they lose the advantage that the display layout could be improved. In fact, we sometimes observed the same phenomena in our experimental studies mentioned in section 4.

Assume that they have the meeting records in the form of Figure 6. In contrast, the participants can refer to the records with the speech act "Explaining Details" and the domain-specific keyword "renumbering" as a retrieval key, and can find the reason as shown in Figure7 (b). As a result, they did not exclude the renumbering function but find the alternative solution, attaching two kinds of number-tags on the nodes.

The point of this example is that we can use speech act information and domain-specific words to get decision rationale. If the size of documents and the meeting time increases, it is difficult for us to remember all of the elicitation processes or to retrieve useful information efficiently. So extracting and holding the elicitation processes with speech act information is one of the effective techniques to solve this problem. That is to say, it can be used as a retrieval key to get the useful information.

4 Experiments for Assessment

In this section, we introduce the experiments on actual meetings of two projects and their analytic results. Furthermore we assess our technique mentioned in the previous section based on the results.

4.1 Experimental Meetings

We applied our technique to two development projects shown in Table 1.

The project#1 is for enhancing the hyper media tool that have been used. An analyst has already collected users' requirements by questionnaires, so the first meeting began with his introducing the users' requirements and his solutions to the requirements. The analyst and two representative users had four meetings and came to final agreement in the fourth meeting. In the project#2, the developers would consider how to support the procedure of a program chairperson of the international academic conference such as ICRE'96. The customer, a program chair, explained his tasks and problems in the first meeting. The analyst presented his solutions to the problems and they discussed the analyst's solutions from the second to the final meetings. Each meeting of projects #1 and #2 was from about 40 minutes to an hour and a half long. All the utterances in the meetings were recorded. Just before a meeting, some documents such as minutes and materials were distributed if any.

In the section 4.3, on account of space, we will introduce the analytic result of the first meeting of the

No.	Speaker	Speech Act	Utterance Text
1	User	Explaining Details	We often have a series of meetings, and we always refer to a specific topic in the agenda or minutes with its number. For example, I say "I would like to discuss again, topic number 9 in the agenda of the previous meeting...".
2	Analyst	Proposing	OK. So we should not change the topic-node numbers until the series of the meetings ends.
3	Analyst	Explaining Details	But we have decided that the nodes in any agenda or minutes should be renumbered when some nodes were deleted or inserted.
4	Analyst	Proposing	Shall we exclude the renumbering function?
5	User	Agreement	Hum, I cannot remember the reason why we adopted the renumbering function, so we don't need it.

(a) Utterance Example #1

No.	Speaker	Speech Act	Utterance Text
		⋮	
4	Analyst	Proposing	Shall we exclude the renumbering function?
5	User	Proposing	Hum, wait a moment, I look for the meeting record related to the topic "renumbering". (referring to the structure shown in Figure 6)
6	User	Explaining Details	Because of improving the display layout, we have decided that the nodes should be renumbered.
7	Analyst	Proposal	So, we can attach the two kinds of number-tags to the topics, one comes from the order of occurrences. We never change this numbering. The other one may be renumbered for improving the display layout, it looks like section, subsection and subsubsection numbers.
8	User	Agreement	Good idea.

(b) Utterance Example #2

Figure 7: Example of Subsequent Utterances

Table 1: Experimental Projects

	Project#1	Project#2
System	Hyper Media Tool	Supporting Program Committee (PC) Chairs
Participants	Analyst, Two Users	Customer (PC Chair), Analyst, Designer
Meeting#1	Analyst's Presentation	Customer's Explanation
#2	Analyst's Presentation	Analyst's Presentation
#3	Analyst's Presentation	Analyst's Presentation
#4	Analyst's Presentation, Agreement	Analyst's Presentation, Agreement

project #1, which was already introduced in the previous section.

4.2 Analytic and Assessment Procedure

Figure 8 shows the flow of our analytic studies. The assessment point is whether our technique could "correctly" identify speech acts of the utterances. The word "correctly" may be ambiguous because there might be no correct answers about what speech act an utterance has. In other words, the interpretation of an utterance together with its context may be ambiguous in some case and its ambiguity leads to impossibility of identifying a speech act. To decide "correct" speech acts of utterances, we emphasize objectivity. In the case that almost persons have the same opinion, we consider that the opinion is "correct". Thus we have two sub flows of our analytic studies — one is for deciding "correct" speech acts objectively and the other one is for applying our technique. As shown in the left sub flow, we have three workers to decide the "correct" speech acts. The right sub flow illustrates the flow of our technique, and it is performed by a worker different from the three workers. We compare the results of applying our technique with the "correct" results by the three workers. The details will be mentioned together with its result in the next sub section.

4.3 Analytic Results

The first meeting had 289 utterances of the project #1. After making utterance text, our worker decided the keywords (including domain-specific keywords), speech acts and typical patterns of speech act transitions. He partitioned the utterance text into segments by using his decided domain-specific words. A part of his selected domain-specific keywords and their occurrences are shown in Figure 9 in addition to Figure 3. For extracting a segment, he used not only a single word but also a set of the words that have semantical relationship to each other. In particular, he often used pairs of a verb and a noun which co-occur in the utterance records. In the case of Figure 3, the word "sort" occurred with the word "number" in the utterance record. The reason is that the renumbering function contained "sorting nodes by numbers" in the system. So he used these words to extract the segment whose topic was concerned with the renumbering function. Each row in Figure 9 shows a set of the words that have semantical relationship. The occurrence times is measured in one minute interval. For example, the words "Parent", "Child" or "Relation" was uttered at four times in the interval between 6 and 7 minutes after the meeting started. He could have topics from the utterance records by the domain-specific keywords of Figures 3 and 9, and finally composed the specification document having seven items.

Next, the other three workers labelled the segmented utterance text with speech acts independently. The speech acts they used have been already introduced in the previous section, i.e. "proposing", "explaining details (incl. rationale)", "questioning", "answering yes", "answering no", "answering others", "agreement" and "disagreement". These speech acts were decided by the worker who partitioned the utterance record in the previous step, i.e. different from the workers who labelled the utterances. If a worker could not find a suitable speech act of an utterance, he did not label it. If the labelling of three workers split, we decided a speech act by majority. It means that a speech act of an utterance can be identified only if more than one worker label the utterance with the same speech act. The majority principle allows us to have stable and objective results on labelling of speech acts. As a result, we could have the result that 73% of the utterances, i.e. 211 of 289, have "correct" speech acts.

Table 2: Keywords for Recognizing "Proposing"

I think, I recommend, I propose, Shall I, Then, Or, Otherwise, Well, No other way

Table 3: Identifying Speech Act by Keywords

Speech Act	Total	Correct	Error	(%)
Proposing	70	40	13	57
Agreement	18	11	1	61
Disagreement	26	19	9	73
Explaining Details	17	14	4	82
Questioning	43	25	11	58
Answering Yes	9	8	0	89
Answering No	2	1	0	50
Answering Others	15	9	1	60
Not Classified	11	–	–	–
Total	211	131	52	62

(%) = Correct ÷ Total

The keywords used for recognizing "proposing" utterances are listed in Table 2. Conversely speaking, "proposing" utterances often contain these words in Japanese. Again, note that the words are direct translations from Japanese to English and one-to-one translation is impossible in some case. In fact, the original Japanese words that we used were 24, and 14 of the 24 words appears at the head of utterances and the remainder at their tails[2]. Variations of the words and the phrases in English, e.g "I would like to recommend" and "We think" etc., are omitted in the table on account of space.

Table 3 shows the comparative results of identifying speech acts by the keywords with three workers' labelling. As the table shows, 70 utterances of 211 utterances, which are successfully labelled by the three workers, were decided to have the speech act "proposing". 40 utterances of the 70s contained the keywords listed in Table 2, while the utterances which were not "proposing" despite containing the keywords were 13. The comparative result tell us that, given a speech act, we can detect correctly about 60% of the utterances having the speech act by means of the keyword search.

Figure 10 shows the transition patterns which were

[2]In Japanese sentences, words in their heads and in their tails, i.e. the first words and the last words have significant effect on their meaning.

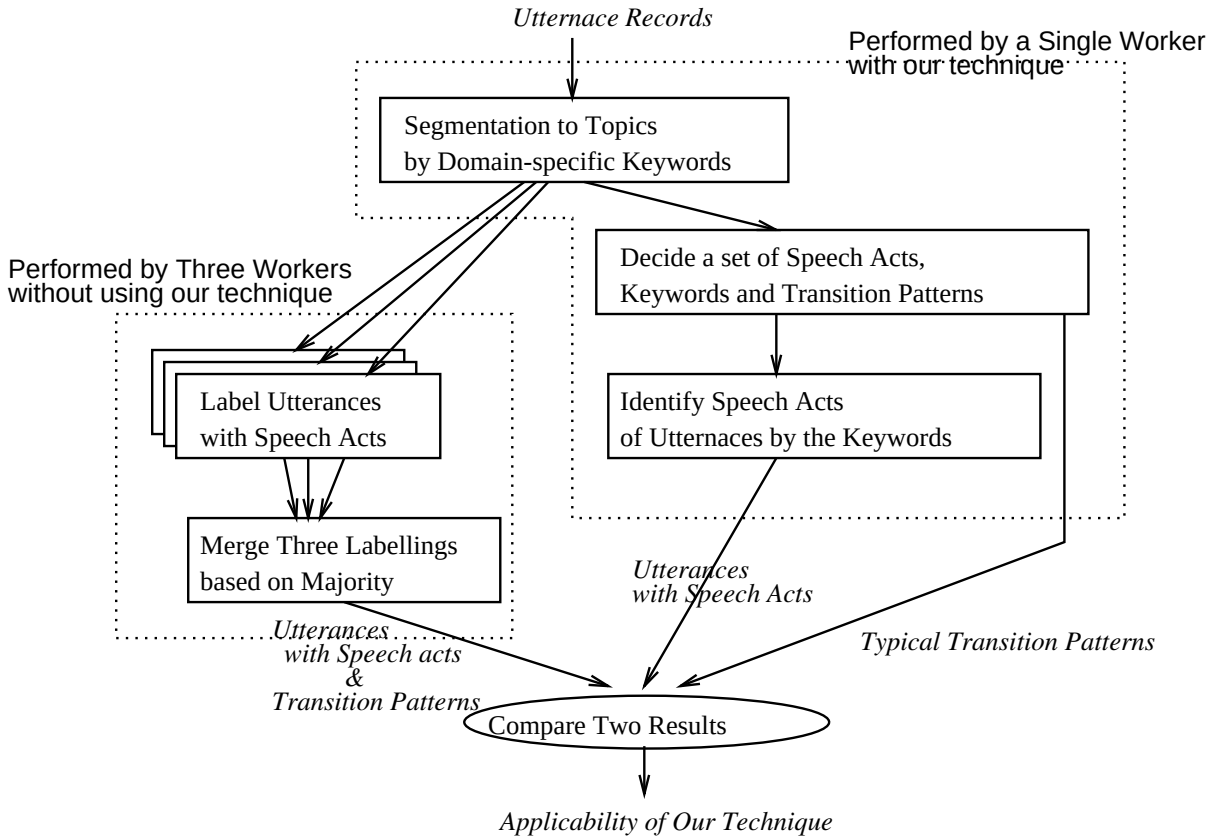

Figure 8: Flow of Analytic Studies

Occurrence Times (one minute interval) 10 20 30	total	Domain-specific Keywords
0000004002 0000100000 0500003643 0004073	42	Parent, Child, Relation
0000000000 1100031581 2050000010 0000000	28	Resize, Magnify, Reduce
0000000001 6873775200 1000000000 0000000	47	gs, Ghostscript, ps, Postscript, Figure, Drawing, Screen
0001000100 0000000343 1300000000 0000000	16	Font, Character, Dot
0000102000 1000000200 0100000100 0003667	32	Pack, Connect, Cut, Delete, Cancel, Dismiss
0000000000 0000221001 1200000001 1110000	13	Structure, Tree, Hierarchy

Figure 9: Occurrences of Domain-specific Keywords

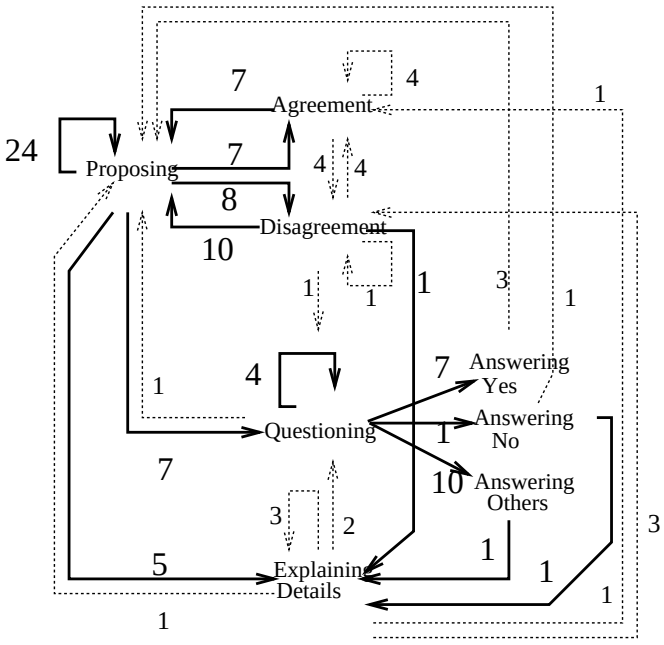

Figure 10: Extracted Transitions of Speech Acts

extracted from the utterance records of the meeting #1. Each number on a transition (arrow) in the figure expresses how many times the transition occurred. The transitions drawn with dotted lines are not the typical ones and they did not appear in Figure 5. As the transitions where pairs of the speech acts could be decided were 122 in total, we can see that 84 % of the transitions follow the typical patterns of Figure 5.

From the analyst's presentation phases in the other meetings listed up in Table 1, we have got the similar results, i.e. about 60 % of speech acts could be detected by the keyword set and 70 or 80 % of the transitions followed the typical pattern. We have an interesting finding about the occurrence times of the speech acts. As the meetings were held repeatedly, the occurrences of the speech act "explaining details" increased while "proposing" decreased. In the third meeting, we had 15 "proposing" and 63 "explaining details" in the 190 utterances while 70 "proposing" and 17 "explaining details" in the first meeting. The analyst tended to concentrate not on proposing new solutions to the requirements but on getting users' agreements by explaining the details in the meetings close to the last meeting.

5 Discussion

In the following part of this section, we discuss the limitations and the directions of future work which is necessary to put our technique into practice.

5.1 Scope of Our Technique

Our goal in this paper is to introducing the technique to transform the meeting transcriptions into requirements elicitation processes. Extracting and holding the elicitation process is very useful to manage the requirements changes. And when the participants discuss the requirements to be decided, for example, they check what is already decided or not, it is very useful to refer to the records of the elicitation processes which were performed before.

We should note that this paper presents just extracting elicitation processes not requirement elicitation itself yet. It means that the paper has not yet referred to the technique to elicit from the record of the elicitation processes the complete requirements which all the participants agreed and came to conclusions. For example, we do not mention the technique or the algorithm to detect automatically conflict requirements and to extract the topics to be debated from the elicitation processes. The studies discussed in [3] and [2] suggest this kind of techniques and our technique may be applicable in a preceding stage of requirements elicitation based on them.

5.2 Keywords

It is possible to collect domain-specific keywords to a thesaurus as domain knowledge for each problem domain. These lexical resources are also useful after requirements analysis phase. In this paper, we have not discussed the advantages of the combination of multiple keywords so much. However focusing on the co-occurrences of the multiple words increases the quality of the segmentation of utterance records. In Figure 3, we simply used the combination of a verb and a noun. We often use demonstrative pronouns such as "it", "they", "that", etc. in conversation. These words can help us identify the referential relationship between utterances. It is one of our future work how to deal with them.

5.3 Speech Acts

A set of speech acts with which utterances are labelled depend on the phases of meetings. In this paper, we considered analysts' presentation phase alone. We should have many case studies to extract a stable set of speech acts and their transition patterns. The keywords for identifying speech acts also depend on speakers. Thus, when identifying its speech act, we should consider who gives an utterance.

In speech act transition patterns, we only considered a pair of the temporally adjacent utterances. In fact, we found that some utterances which were temporally separated had referential relationship in the utterance records. For example, in the case that several participants answered yes to a question, its sequence of speech acts was "questioning – answering yes – $\cdots$ – answering yes". The last "answering yes" did not appear immediately after the "questioning" and this may make it difficult to identify the relationship among them. Thus we should consider not only an utterance immediately after the utterance but also several utterances after it. Exploring co-occurrences of domain-specific words help us identify referential relationship among the utterances which are temporally separated.

We never think that all of the utterances can have a unique speech act, so it is ambiguous for determining speech act of an utterance. In fact, we had much experiences that an utterance had several speech acts. If the speech acts of the utterance is really ambiguous, it is really ambiguous to human beings. In this case, it is

useless to force computers to determine which speech act it has but the case can suggest to its speaker that he or she should speak it again more clearly. Our technique may be able to detect which utterances have ambiguity. One problem is when this ambiguity detection and clarification should be done. If it is performed online, i.e. during the meeting, it may interfere with the natural flow of the conversation. Usually, we have a series of the meetings until getting agreement. We think that the ambiguity detection would be done among the meetings to get the natural conversation flow. If the analysts detect the ambiguous utterances and they cannot solve them, they pick up their clarification as agenda in the next meeting.

As some linguists pointed out[8], there are limitations of speech act theory to model human communication. Speech act theory could not handle automatic requirements elicitation or represent linguistic phenomena, e.g. detecting contradictions appearing in requirements elicitation, although we could make elaborated speech act theory. We do not think that speech act theory can formalize all phenomena in the elicitation process. It is helpful just for the participants to retrieve information they want by using speech acts as a kind of retrieving keys. To elicit requirements from the records of the elicitation processes, including how to discuss the topics in the meetings, is essentially human task. The structure of the records can just support the activities of this task. As shown in section 3.3, our participants could elicit alternative solution with the aids of the structured process.

5.4 Speech Processing

The task to write down the utterances to text needs much human efforts. If it could be automatically or semi-automatically performed by computer, we could reduce the difficulty of our tasks. Currently complete speech recognition on computers, in particular recognition of non-restricted conversation, are technically still impossible. However word spotting technique to extract specific words from conversation has paid off very much. Since our technique is based on focusing specific keywords, automatic processing on utterances has bright possibility by using word spotting technique.

Acknowledgements

The authors would like to thank Dr. John Dobson for insightful comments to the earlier version of this paper. They would also like to thank Prof. K. Ochimizu of Japan Advanced Institute of Science and Technology, Hokuriku for discussions about this paper.

References

[1] E. Auramäki, E. Lehtinen, and L. Lyytinen. A Speech-Act-Based Office Modeling Approach. *ACM Trans. on Office Information Systems*, Vol. 6, No. 2, pp. 126–152, 1988.

[2] B. Boehm, P. Bose, E. Horowitz, and M. Lee. Software requirements Negotiation and Renegotiation Aids : A Theory-W Based Approach. In *Proc. of 17th ICSE*, pp. 243 – 253, 1995.

[3] L. Chung and B. Nixon. Dealing with Non-Functional Requirements: Three Experimental Studies of a Process-Oriented Approach. In *Proc. of ICSE17*, pp. 25–37, 1995.

[4] T. DeMarco. *Structured Analysis and System Specification.* Yourdon Press, 1978.

[5] A. Finkelstein. Requirements Engineering: A Review and Research Agenda. In *Proc. of 1st Asia-Pacific Software Engineering Conference*, pp. 10–19, 1994.

[6] J. Goguen and C. Linde. Techniques for Requirements Elicitation. In *Proc. of First IEEE International Symposium on Requirements Engineering*, pp. 152–164, 1993.

[7] O. Gotel and A. Finkelstein. Contribution Structures. In *Proc. of Second IEEE International Symposium on Requirements Engineering*, pp. 100–107, 1995.

[8] J. (English Translation by T. McCarthy) Habermas. *The Theory of Communicative Action.* Beacon Press, 1984.

[9] P. Johannesson. Representation and Communication in Information Systems — A Speech Act Based Approach. In *Lecture Notes in Computer Science (CAiSE'94)*, pp. 200–213. Springer-Verlag, 1994.

[10] T.G. Lewis. *CASE : Computer-Aided Software Engineering.* Van Nostrand Reinhold, 1991.

[11] C. Nellborn and P. Holm. Capturing Information Systems Requirements Through Enterprise and Speech Act Modelling. In *Lecture Notes in Computer Science (CAiSE'94)*, pp. 172–185. Springer-Verlag, 1994.

[12] C. Potts, K. Takahashi, and A. Anton. Inquiry-Based Requirements Analysis. *IEEE Software*, Vol. 11, No. 2, pp. 21–32, 1994.

[13] J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy, and W. Lonrensen. *Object-Oriented Modeling and Design.* Prentice-Hall, 1991.

[14] J.R. Searle. *Speech Acts : An Essay in the Philosophy of Language.* Cambridge Univ. Press, 1969.

[15] S. Shlaer and S.J. Mellor. An Object-Oriented Approach to Domain Analysis. *ACM SIGSOFT Software Engineering Notes*, Vol. 14, No. 5, pp. 66–77, 1989.

[16] T. Winograd. Where the action is. *BYTE*, Vol. 13, No. 13, pp. 256–260, 1988.