ICOT Technical Report: TR-443

TR-443

Amount of information of a Sentence Interpretation against a Knowledge Base

by

K. Sumita. T. Ukita and S. Amano(Toshiba)

December. 1988

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

Title:

Amount of information of a sentence interpretation aginst a knowledge base*

Author:
Kazuo SUMITA, Teruhiko UKITA and Shin-ya AMANO

Toshiba Corp. R & D Center

Address:
Toshiba Corp. R & D Center, Information Systems Lab.
Komukai-Toshiba-cho 1, Saiwaiku, Kawasaki 210, Japan

Area:
(c1) Natural language

ABSTRACT

This paper describes the amount of information of a sentence interpretation, which is used as a measure for the disambiguation process in a natural language understanding system. This measure is defined, based on a hearer's model with knowledge base composed of proposition set and inference rule set. Selecting the most informative interpretation by this measure is reasonable in the sence that communication is an act whereby messages are transmitted with the least effort. This formalization is applied to a practical procedure for anaphoric ambiguity resolution, which is constructed as a part of a question-answering system. Furthermore, a conversation experiment was carried out, and it was found that ninety-three percent of referents corresponding to anaphoric indicators could be correctly determined.

* This work was supported by ICOT (Insitute for New Generation Computer Technology), and was carried out as part of the FGCS (Fifth Generation Computer Systems) project in Japan.

1. INTRODUCTION

To build a natural language understanding system, it is necessary to manage many sentences which don't contain enough information to transmit their meanings. Anaphoric references, ellipses, fragmental input, etc. are examples of such discourse phenomena. In the process of recovering insufficient expression, ambiguity problems arise. Criterion is required for ambiguity resolution; a natural language understanding system should select a plausible one out of several possible interpretation candidates using the criterion.

In conversation, a speaker would transmit his/her message with less effort. Grice's conversational maxim of quantity [Grice 75] supports this idea. Thus, a hearer should understand the message on the basis of this maxim to get the most informative interpretation. However, it is necessary to give a quantative measure for practical procedures, since this maxim is just a qualitative principle.

The amount of information on a knowledge representation was discussed by Ohsuga [Ohsuga 84]. This measure is a criterion appropriate for context understanding, which is a process of disambiguating the hearer's knowledge. The amount of information of a predicate on ***many sorted logic*** was defined in his research. However, knowledge representation, including inference rules, was't considerd, while the role of causality relations among events is important in a context upderstanding process. Thus, it is necessary to extend the application range of this measure to a knowledge base with inference rules.

This paper defines the amount of information of a sentence interpretation, which is based on a hearer's comprehension model including knowledge about causality relations among events. In the following sections, a comprehension model of hearer and amount of information of a sentence interpretation are discussed. This measure is applied to anaphoric ambiguity resolution. Finally, a conversation experiment is carried out, using a question-answering system with a knowledge processing sub-system.

2. HEARER'S COMPREHENSION MODEL

Figure 1 shows a hearer's comprehension model, which consists of knowledge base and two processes: sentence analysis and preference judgement. At the first process, interpretation candidates against a new input sentence are produced. At the second process, the most plausible one is determined, according to the knowledge base, to which the obtained interpretation is added as new knowledge. In general, the knowledge base includes linguistic knowledge, knowledge about facts of the world, causality relations among events, knowledge about the speaker, etc.

In this paper, it is assumed that the hearer's knowledge should be composed of a set of propositions D and a set of inference rules K; in other words, a hearer knows some propositions (D) and some inference rules (K).

A universal set of all propositions Ω is presupposed as follows,

$$\Omega = \{ \omega_n \mid 1 \leqq n \leqq N \}. \qquad (1)$$

D is a proposition set, where each element π_i in D satisfies either $\pi_i = \omega_n$ or $\pi_i = \overline{\omega_n}$ for some element ω_n in Ω. ($\overline{\omega_n}$ denotes negation of ω_n.) Note that D doesn't contain any compositional proposition with conjunction and/or disjunction.

K is a set of inference rules, such that the right-hand side of each element κ_j in K is exactly either an affirmation or a negation of the proposition in Ω. Its left-hand side is a conjunction of some affirmative propositions and/or some negative for those in Ω. For example, D and K are given as follows,

$$D = \{\pi_1, \pi_2, \pi_3\} = \{\omega_1, \omega_4, \overline{\omega_{10}}\}. \quad (2)$$

$$K = \{\kappa_1, \kappa_2\} = \{\omega_3 \wedge \omega_4 \rightarrow \omega_5, \omega_1 \wedge \overline{\omega_2} \rightarrow \omega_3\}. \quad (3)$$

When a sentence is input to the hearer, it is analyzed into sevral interpretations. Thus, it will be assumed that a proposition ν is a sentence interpretation, which is added to the knowledge base. Then, some new propositions are derivable from ν, D and K. Let A be a set of the derived propostions. After D is replaced by a union D' of D, $\{\nu\}$ and A, other sentences are accepted repeatedly; D' is a new proposition set which the hearer knows. That is

$$\nu, D, K \vdash A. \quad (4)$$

$$D' = D \cup \{\nu\} \cup A. \quad (5)$$

It is assumed that the proposition set D in the hearer's knowledge is consistent; if $\pi_i \in D$, then $\overline{\pi_i} \notin D$.

3. AMOUNT OF INFORMATION IN AN INTERPRETATION

How to calculate the amount of information of an interpretation is discussed, based on the hearer's comprehension model described above.

3.1 Definition of amount of information in an interpretation

Figure 2 shows the states of the hearer's knowledge, before and after he has recieved a statement that the button was pushed. Before he has heard the statement, he didn't know whether the button was pushed or not; there were two possibilities in his knowldge state. On the other hand, after he has heard the statement, only one possibility remains. The function of input sentences is to reduce the possible states of a hearer's knowledge.

To formalize the above situation, let's consider possible combinations of all propositions. When truth values for all propositions in the universal set Ω are determined in the hearer's knowledge D, let us regard the sequence of these truth values of propositions as a state of the hearer's knowledge. For instance, if all propositions exept ω_1 are true, then the state can be written as follows,

$$< \overline{\omega_1}, \omega_2, \omega_3, \cdots, \omega_N >. \quad (6)$$

S_Ω(D,K) is defined to be a set of states which are consistent with D and K .* From now on, the subscript Ω in S_Ω(D,K) will be omitted for simplifying descriptions, because it was assumed that Ω is fixed.

For example, assume that D=ϕ and K=ϕ . Since D doesn't contradict all possibilities of combinations regarding affirmative propositions and negative propositions, this set of states is given as follows,

$$
\begin{aligned}
&S(\phi, \phi) \\
&\quad = \{< \omega_1, \omega_2, \omega_3, \cdots, \omega_N >, \\
&\qquad < \overline{\omega_1}, \omega_2, \omega_3, \cdots, \omega_N >, \\
&\qquad < \omega_1, \overline{\omega_2}, \omega_3, \cdots, \omega_N >, \\
&\qquad \qquad \cdots \\
&\qquad < \overline{\omega_1}, \overline{\omega_2}, \overline{\omega_3}, \cdots, \overline{\omega_N} >\}. \qquad (7)
\end{aligned}
$$

Every state in S(D,K) is mutually exclusive. Also, it is unknown which states hold true. We make the following assumption.

Assumption 1:
A-priori probabilities for every state in S(D,K) are equal to each other.

Under this presupposition, E(D,K), entropy of S(D,K), is obtained:

$$
\begin{aligned}
E(D,K) &= - \sum^{|S(D,K)|} |S(D,K)|^{-1} \log_2 |S(D,K)|^{-1} \\
&= \log_2 |S(D,K)|, \qquad (8)
\end{aligned}
$$

where |S(D,K)| means the number of elements in S(D,K). Assume that a new proposition ν is given, and that D is replaced by D' obtained from Eq. (5). The amount of information in this proposition ν , I(ν ,D,K), is defined as follows,

$$
\begin{aligned}
I(\nu, D, K) &= E(D,K) - E(D',K) \\
&= \log_2 (|S(D,K)| \,/\, |S(D',K)|). \qquad (9)
\end{aligned}
$$

The above Eq.(9) indicates that the amount of information in a proposition is calculated by the number of elements in the set of states.

In the case of knowledge without inference rules, Ohsuga discussed a similar calculation process for amount of information of a proposition at quantative consideration for knowledge representation [Ohsuga 84]. Here, brief summary of the consideration is presented for subsequent explanation. Since his theory does not contain inference rules; K=ϕ in our notations. For the case D=ϕ , it is clear that E(ϕ , ϕ)=N, since |S(ϕ , ϕ)|= 2^N from Eq.(7). For instance, assume that a new proposition against the proposition set

* If a state is consistent with every proposition in D, then we say that the state is consistent with D, and if a state is consistent with every compositional proposition equivalent to the inference rule in K, then we say that the state is consistent with K.

D (= ϕ) is ω_1. Then, D'={ ω_1 }. Since S({ ω_1 }, ϕ) is a set of a states which are consistent with ω_1, and are extracted from S(ϕ , ϕ), clearly |S({ ω_1 }, ϕ)| is half of |S(ϕ , ϕ)|. Thus, |S({ ω_1 }, ϕ)|= 2^{N-1}, and I({ ω_1 }, ϕ)=1. In general, if $\nu \notin$ D and K= ϕ , then D'= D $\cup$ { ν }, and |S(D',K)| is half of |S(D,K)|. Therefore, the amount of information of a proposition ν is 1.

3.2 Amount of information in the case of knowledge with inference rules

Let's discuss how to calculate the amount of information, when knowledge consists of propostions and inference rules. Some inference rules are dependent on each other; sets of inference rules which share common propositions can be assumed. Let a member of these sets be written as Γ h. It is evident that K is divided into each Γ h completely. Also, proposition sets Δ h, which correspond to Γ h, are classified; Δ h is a set of propositions which appear on inference rules in Γ h. Figure 3 shows this relation between Γ h and Δ h. When either ν or $\overline{\nu}$ exists in the set of propositions constructing some inference rule κ j in K, we say that the proposition ν relates to inference rule κ j. Let a set of all combinations for propositions consistent with D and K, in Δ h, be C(Δ h,D, Γ h). Then, the following theorem is obtained.

Theorem 1:

If a proposition ν relates to some inference rule, then indicate a set Γ h, to which the inference rule belongs. The amount of information of ν can be calculated by combinations of all propositions in Δ h corresponding to Γ h. That is as follows,

$$I(\nu, D, K) = \log_2 (|C(\Delta h, D, \Gamma h)| / |C(\Delta h, D', \Delta h)|). \qquad (10)$$

Proof.

No two proposition sets Δ_i and Δ_j for i $\neq$ j, share any common proposition, because of its definition. Also, let the number of conjunctive combinations of those which don't belong to any Δ i be C0 (C0= 2^{N0-L}, where N0 is the number of propositions which don't belong to any Δ i, and L is the number of propositions, out of those propositions, whose truth values are determined in D already.). Then, the number of elements in S(D,K) is obtained as the product of each number of the combinations C(Δ h,D, Γ h);

$$|S(D,K)| = C0 \times \prod_h |C(\Delta h, D, \Gamma h)|. \qquad (11)$$

Substituting (11) into (9) gives (10). Q.E.D.

The amount of information of any proposition ν can be calculated in accordance with the above theorem, which is given by

$$I(\nu, D, K) = \begin{cases} \log_2 (|C(\Delta h, D, \Gamma h)| / |C(\Delta h, D', \Gamma h)|) \\ \qquad (\nu \text{ relates to some } \kappa \text{ j in } \Gamma h) \\ 1 \qquad (\nu \text{ doesn't relate to any } \kappa \text{ j in K}) \\ 0 \qquad (\nu \in D). \end{cases} \qquad (12)$$

To illustrate this process accurately, let's consider an example. Consider the case where D={f}, K={a^b → c, c^$\bar{d}$ → e, f^g → h, g → $\bar{q}$}, Ω ={a,b,c,d,e,f,g,h,p,q,r} and ν =$\bar{e}$; let a^b → c, c^$\bar{d}$ → e, f^g → h and g → q be κ_1, κ_2, κ_3 and κ_4, respectively. Then, K can be divided into Γ_1= { κ_1, κ_2} and Γ_2= { κ_3, κ_4}, and Δ_1={a,b,c,d,e} and Δ_2={f,g,h, q} can be extracted from Ω , which correspond to Γ_1 and Γ_2, respectively. On the other hand, propositions "r" and "p" don't belong to either Δ_1 or Δ_2. Since the negation of "$\bar{e}$" is equal to the right-hand side of κ_2, "e" relates to κ_2 in Γ_1. Thus, combinations of all propositions in Δ_1 corresponding to Γ_1 should be taken into account. Figure 4 depicts these combinations; there are 32 combinations in this matrix. At the first step, combinations conflicting with either κ_1 or κ_2 are removed, which are marked with "×" in Fig.4, after which 24 combinations remain. In this case, since individuals in D do not exist in Δ_1, there are no combinations which conflict with D. At the second step, combinations conflicting with a proposition ν =$\bar{e}$ are removed, which are marked with "✓" in Figure 4. It is easy to see |C(Δ_1,D', Γ_1)|=10. In consequence, the amount of information on "$\bar{e}$" is obtained; I($\bar{e}$,D,K)=$\log_2 2.4$.

When individual inference rules do not share common propositions

If the inference rule set K satifies the following condition, then it is not necessary to pay attention to combinations of propositions.

Condition 1:

For each κ_i and κ_j, where i ≠ j, there is no common proposition which is shared by both κ_i and κ_j.

Under this condition, an inference rule κ_j is separated from others. Therefore, the amount of information of ν depends on only κ_j, which ν relates to. As a result, the following equation is obtained.

$$I(\nu,D,K) = \begin{cases} \log_2 \{ (2^{m_j - k_j} - \sigma_j(D)) / (2^{m_j - k_j - 1} - \sigma_j(D')) \} \\ \quad (\nu \text{ relates to some } \kappa_j) \\ 1 \quad (\nu \text{ doesn't relate to any } \kappa_j) \\ 0 \quad (\nu \in D), \end{cases} \qquad (13)$$

where m_j is the number of propositions composing κ_j, and k_j is the number of propositions, out of those propositions, whose truth values are known in D already. And σ_j(D) is a function such as gives 1 or 0; if every proposition, in D, which relates to κ_j belongs to a minimum set of propositions which contradicts κ_j, then σ_j(D)=1, otherwise, σ_j(D)=0.

4. ANAPHORIC AMBIGUITY RESOLUTION

Formalization for an amount of information described above is applied to the problem of resolving anaphoric ambiguity. An informative interpretation candidate is selected so that it could give the largest amount of information, under the condition that individual inference rules do not share common propositions.

The anaphora and ellipsis have been analyzed and managed from various points of view (e.g., [Webber 80]), such as syntactic and semantic approach, inferential approach, etc. Sidner proposed a ***bootstrapping*** procedure using ***focus***, which serves as a primary antecedent, as long as it doesn't violate semantic constraints [Sidner 83]. Her procedure can't select the most preferable candidate, when there are plural consistent candidates. Hobbs tried to resolve anaphoric ambiguity by relations between events [Hobbs 79]. Since his approach is not based on any quantative comprehension model, it lacks in objectivity.

4.1 Problem

Figure 5 shows an input sentence in a consultation dialogue on VTR (video tape recorder) operation. In this example, an interpretation candidate of "it=VTR" is preferable to that of "it=cassette-tape" or "it=playback-button", though they satisfy the semantic constraint of "work". This preference comes from the fact that the hearer has the following knowledge:

If someone inserts a casette tape into a VTR,
and pushes the playback button,
then the VTR will work.

Using such causality relations, an informative interpretation should be selected, since the ambiguity whcich arose in Fig.5 can't be resolved by syntactic and/or semantic constraints alone.

4.2 Procedure

A practical procedure to resolve anaphora is described, which is based on amount of information on a sentence interpretation. The procedure consists of three major steps; anaphora detection, referent candidate extraction and preference judgement.

The procedure has been incorporated as a part of an experimental question-answering system, whose task is guidance of VTR operation [Ukita 88]. The system uses a knowledge representation system [Kinoshita 88], which represents an object and an event as a ***schema*** in a similar way as ***units*** in KRL [Bobrow 77] and causality relations among events by if-then rules. This Q/A system has been developed on PSI-II (PSI-II is a computer system developed by ICOT), and owns 1000 lexicons, 300 schemata and 30 rules currently.

Every input sentence in Japanese is analyzed syntactically and semantically, and is represented as a dependence structure of case-frames. For each case-frame, the following procedure is executed.

<u>Anaphora detection</u> As the first step of anaphora disambiguation, the procedure detects anaphoric indicators, such as pronouns, definite nouns, omitted obligatory cases of the predicate in the case-frame being processed, and nouns themselves. Extracted

indicators are managed as tentative instance schemata.

Referent candidate extraction For every anaphoric indicator, the procedure searches for referent candidates corresponding to the indicator. Instance schemata, which belong to the same class of the indicator, those which belong to a sub-class of indicator's class, and those which belong to a class linked to the indicator by has-part relation, are extracted as candidates. On the other hand, if an anaphoric indicator refers to some event, the procedure searches for preceding events in the conversation history.

Preference judgement When plural interpretation candidates remain, the procedure tries to decide the most informative interpretation by inspecting if-then rules and preceding events. An interpretation, which maximizes the amount of its information calculated by Eq.(13), is selected as a plausible one.

After the procedure selects interpretations for all case-frames, it accepts next setences. Figure 6 shows a process exapmle for "Though I pushed the VTR playback button, it's not working."

5. AN EXPERIMENT

A conversation experiment has been carried out using the experimental Q/A system with the procedure for anaphora resolution described above. Nine dialogues, such as shown in Fig.7, were given to four persons, and sentences which can follow these dialogues were collected. Among these sentences, 113 sentences which the system can analyze syntactically, have been used for the experiment. There were 174 clauses in the sentences. Table 1 shows how many referents corresponding to anaphoric indidators (total 204) can be found correctly.

Fourteen errors were due to 5 causes; (a) exact objects were too far (e.g., more than 10 objects apart) from anaphoric indicators to be found as candidates (3 times), (b) no event candidates were extracted at the step of referent candidate extraction (3 times), (c) there were no schemata linking to lexical words in the knowldge base (4 times), (d) schemata corresponding to predicates didn't have correct slots (3 times), (e) anaphoric indicators were used with interrogative pronouns (1 time; "Where shall I insert it?"). As to first cause, if the procedure would search more than 10 preceding objects, it could find correct answers. Also, the former 4 causes come from the step of referent candidate extraction.

The procedure produced plural interpretation candidates 47 times, and it failed to select correct candidates 6 times. These errors, which correspond to above (a) and (b), were due to insufficiency of the candidate extraction step. On the other hand, all interpretations that related to some if-then rule were selected correctly, and it was 32 times. As a result, 91 percent of interpretations were decided exactly in total.

If the procedure would omit the preference judgement step, it could decide only 80 percent of interpretations correctly, which are composed of most nearest objects satisfying semantic constraints. Therefore, this step increased the accuracy on the interpretation decision by around 10 percent.

6. CONCLUSION

The amount of information of a sentence interpretation, as the measure for disambiguation process of a natural language understanding system, is described. This measure is defined on the basis of on a hearer's comprehension model including knowledge about causality relations among events as well as propositions. A theory of calculating the amount of information of a possible interpretation is described. Furthermore, this measure is applied to the procedure for anaphora disambiguation, which is constructed as a part of a question-answering system. Finally, a conversation experiment was carried out using this system. Ninety-three percent of referents can be detected correctly. Even though many sentences are not used in the experiment, the effectiveness of the defined measure is confirmed; the accuracy on the interpretation decision rises by around 10 percent.

REFERENCES

[Bobrow 77]
Bobrow, D. and Winograd, T.: KRL:Knowledge Representation Language, Cognitive Science, Vol.1, No.1, 1977.

[Grice 75]
Grice, H.P.: Logic and Conversation, Syntax Semantics, Vol.3, Speech Acts, Seminar Press, pp.41-58, 1975.

[Hobbs 79]
Hobbs, J.R.: Coherence and Coreference, Cognitive Science, Vol.3, pp.67-90, 1979.

[Kinoshita 88]
Kinoshita, S., Sano, Y., Ukita, T., Sumita, K. and Amano, S.: Knowledge Representation and Reasoning for Discourse Understanding, Proc. of the Logic Programming Conf. '88, pp.205-212, 1988 (in Japanese).

[Ohsuga 84]
Ohsuga, S.: A Consideration to Knowledge Representation - an Information Theoretic View, Trans., Information Processing Society Japan, Vol.25, No.4, pp.685-694, 1984 (in Japanese).

[Sidner 83]
Sidner, C.L.: Focusing in Comprehension of Definite Anaphora, M.Brady and R.C.Beriwick (Eds.), Computational Models of Discourse, MIT press, pp.267-330, 1983.

[Ukita 88]
Ukita, T., Sumita, K., Kinoshita, S., Sano, H., Amano, S.: Preference Judgement in Comprehending Conversational Sentences using Multi-paradigm World Knowledge, Proc. of Int. Conf. on Fifth Generation Computer Systems '88, pp.1133-1140, 1988.

[Webber 80]
Webber, B.L.: Syntax beyond the Sentence: Anaphora, in R.J.Spiro, et al. (Eds.), Theoretical Issues in Reading Comprehension, Lawrence Erlbaum Associates, Inc., 1980.

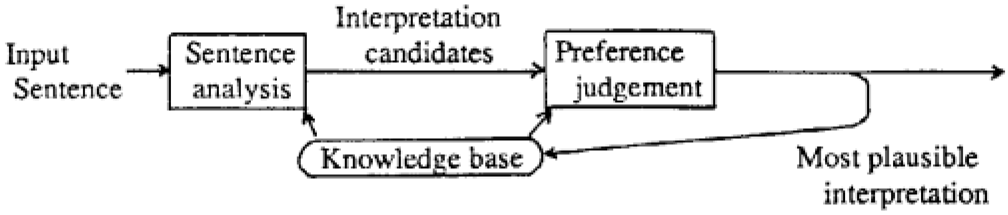

Fig.1 Hearer's comprehension model

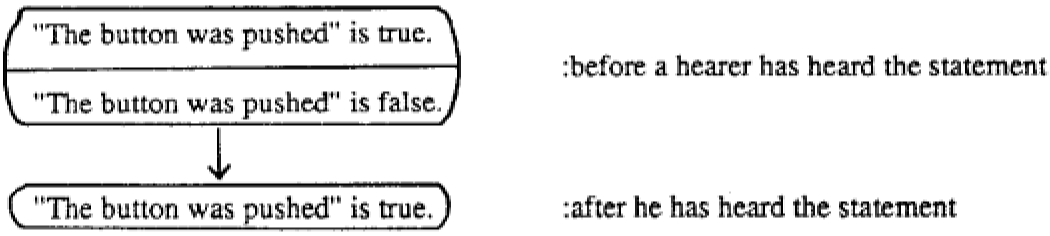

Fig.2 States of hearer's knowledge, before and after he has heard the statement that the button was pushed.

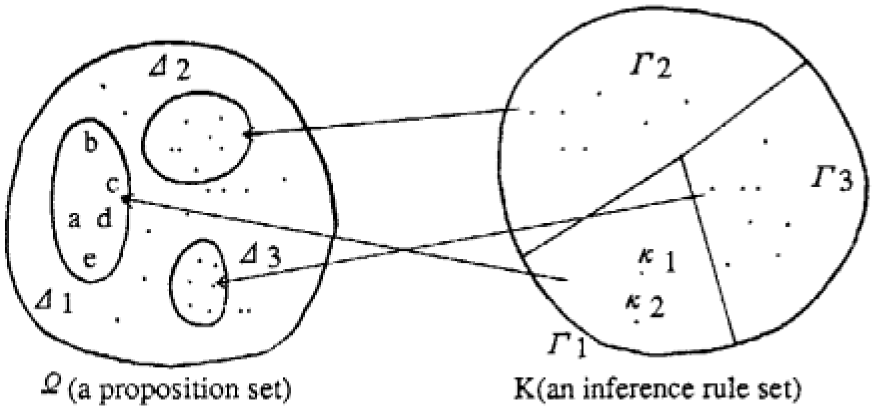

Fig.3 Relation between an inference rule set Γh and a proposition set Δh (κ_1 is $a \wedge b \rightarrow c$, and κ_2 is $c \wedge \bar{d} \rightarrow e$.)

√ a,b,c,d,e	√ $a,\bar{b},c,d,e$	√ $\bar{a},b,c,d,e$	√ $\bar{a},\bar{b},c,d,e$
$a,b,c,d,\bar{e}$	$a,\bar{b},c,d,\bar{e}$	$\bar{a},b,c,d,\bar{e}$	$\bar{a},\bar{b},c,d,\bar{e}$
√ $a,b,c,\bar{d},e$	√ $a,\bar{b},c,\bar{d},e$	√ $\bar{a},b,c,\bar{d},e$	√ $\bar{a},\bar{b},c,\bar{d},e$
× $a,b,c,\bar{d},\bar{e}$	× $a,\bar{b},c,\bar{d},\bar{e}$	× $\bar{a},b,c,\bar{d},\bar{e}$	× $\bar{a},\bar{b},c,\bar{d},\bar{e}$
× $a,b,\bar{c},d,e$	√ $a,\bar{b},\bar{c},d,e$	√ $\bar{a},b,\bar{c},d,e$	√ $\bar{a},\bar{b},\bar{c},d,e$
× $a,b,\bar{c},d,\bar{e}$	$a,\bar{b},\bar{c},d,\bar{e}$	$\bar{a},b,\bar{c},d,\bar{e}$	$\bar{a},\bar{b},\bar{c},d,\bar{e}$
× $a,b,\bar{c},\bar{d},e$	√ $a,\bar{b},\bar{c},\bar{d},e$	√ $\bar{a},b,\bar{c},\bar{d},e$	√ $\bar{a},\bar{b},\bar{c},\bar{d},e$
× $a,b,\bar{c},\bar{d},\bar{e}$	$a,\bar{b},\bar{c},\bar{d},\bar{e}$	$\bar{a},b,\bar{c},\bar{d},\bar{e}$	$\bar{a},\bar{b},\bar{c},\bar{d},\bar{e}$

Fig.4 Combinations of affirmations and negations on all propositions {a,b,c,d,e} (Combinations with × conflict with either $a \wedge b \rightarrow c$ or $c \wedge \bar{d} \rightarrow e$. Combinations with √ conflict with $\bar{e}$.)

Though I inserted a tape into the VTR
and pushed the playback button, it's not working.

Fig.5 A sentence exapmle in a dialogue about VTR operation

Data after the 1st clause ("Though I pushed the VTR playback button,") processed

<instance schema>
schema(instance,vtr#1000, [(class,vtr), (has-part,playback-button#1001)])
schema(instance,playback-button#1001, [(class,playback-button), (part-of,vtr#1000)])
schema(instance,user#1, [(class,user)])
<history>
event(push, yes, [(agent,user#1), (object,playback-button#1001)])

Process for the 2nd clause ("It's not working.")

<referent candidates for "it">
vtr#1000, playback-button#1001
<interpretation candidates for the clause>
(a) event(work, no, [(object,vtr#1000)])
(b) event(work, no, [(object,playback-button#1001)])
<prefernce judgement>
inference rule relating to (a): event(insert, yes, [(object,tape), (goal,VTR)])
event(push, yes, [(object,playback-button)])
→ event(work, yes, [(object,VTR)])
rule to (b): nil
<result of the sentence interpretation>
event(work, no, [(object,vtr#1000)])

Fig.6 A process example of the procedure for disambiguation of anaphoric references
(input sentence:"Though I pushed the VTR playback button, it's not working.")

User: Show me how to playback a tape.
System: Please insert the tape into the VTR, and push the playback button.
(a) An example of given dialogues

May I push the button, before I playback it.
(b) An exmaple of collected sentences which can follow (a)

Fig.7 An exmaple of given dialogues and collected sentences

Table.1 The result of the experiment disambiguating anaphoric references, based on the amount of information of a sentence
(*1 count of referents found correctly,
*2 object : corresponds to nominal entity,
*3 event : corresponds to predicate)

Anaphoric indicator	Referent type	Correct*1 (count)	Total (count)
elipsis	object*2	86	90
	event *3	12	18
pronoun	object	3	4
	event	3	3
noun	object	74	77
definite noun	object	12	12
Total		190	204