TR-I-0344

Transfer of Short Answers from Japanese into English Using a Context-dependent Plan Inference System

John K. Myers

March 12, 1993

Abstract

This paper discusses the phenomenon of short answers and their importance in machine translation, and presents a demonstration of a plan inference system that uses feature structures to recognize and transfer short answers in the middle of a dialog.

Transfer of Short Answers from Japanese into English Using a Context-dependent Plan Inference System

John K. Myers

ATR Interpreting Telephony Research Laboratories
Sanpeidani, Inuidani
Seika-cho, Soraku-gun, Kyoto 619-02, Japan
myers@atr-la.atr.co.jp

Abstract

This paper discusses the phenomenon of short answers and their importance in machine translation, and presents a demonstration of a plan inference system that uses feature structures to recognize and transfer short answers in the middle of a dialog.

In Japanese, short answers to yes/no questions are formed by repeating the verb. English forms short answers by repeating the verbal auxiliary. Thus, although a literal translation of Japanese short answers is semantically well-formed, it is not conventionally acceptable to English native speakers. Accordingly, in an automatic interpretation system, short answers must be recognized and transferred. This recognition is context-dependent.

This process is accomplished by "NP", an assumption-based plan inference system which uses plan schemata represented in a feature structure format. Input to the system can be taken directly from the output of a semantic parser. A short answer is defined as a type of plan action to be recognized. Recognition of a short answer fires a demon that processes the utterance representation and sends it to be transferred. Currently, transfer is accomplished by a rewriting system. The resulting interpreted feature structure can then be sent to a generation system.

In this manner, short answers are transferred. The NP plan inference system is capable of performing similar context-dependent recognition and processing tasks for machine translation.

Abstract

This paper discusses the phenomenon of short answers and their importance in machine translation, and presents a demonstration of a plan inference system that uses feature structures to recognize and transfer short answers in the middle of a dialog.

In Japanese, short answers to yes/no questions are formed by repeating the verb. English forms short answers by repeating the verbal auxiliary. Thus, although a literal translation of Japanese short answers is semantically well-formed, it is not conventionally acceptable to English native speakers. Accordingly, in an automatic interpretation system, short answers must be recognized and transferred. This recognition is context-dependent.

This process is accomplished by "NP", an assumption-based plan inference system which uses plan schemata represented in a feature structure format. Input to the system can be taken directly from the output of a semantic parser. A short answer is defined as a type of plan action to be recognized. Recognition of a short answer fires a demon that processes the utterance representation and sends it to be transferred. Currently, transfer is accomplished by a rewriting system. The resulting interpreted feature structure can then be sent to a generation system.

In this manner, short answers are transferred. The NP plan inference system is capable of performing similar context-dependent recognition and processing tasks for machine translation.

1 Introduction

This work is designed to contribute to an automatic interpreting telephone system which will translate spoken Japanese/English dialogs.

Machine translation of dialogs requires the transfer of culturally-specific context-dependent speech acts. But in order to be able to transfer these speech acts, it is first necessary to recognize and understand them. This is the task of a plan inference system, which must be located between parsing and transfer.

One of the simplest examples of context-dependent speech acts is the short answer to a yes/no question. In English, short answers are formed by answering yes/no and then by repeating the subject and the verbal auxiliaries, which are marked for tense, aspect, positive/negative, etc. However, the verb is not repeated, and the object(s) of the verb is (are) also elided.

Examples:

```
1a. Do you have the announcement?                (Possessive "have".)
1b. Yes, I do.

2a. Have you paid your fee yet?
2b. Yes, I have.                                 (Perfective "have".)

3a. Will you come to the conference?
3b. No, I won't.
```

In Japanese, on the other hand, short answers are formed by an optional yes/no answer, and then by repeating the verb, marked for tense, aspect, positive/negative, etc. The object(s) of the verb is (are) elided, as is the subject/topic.

Examples:

```
4a. Annai-sho wa o-mochi-desu ka? [Announcement TOPIC possess-POLITE-STATIVE QUESTION?]
4b. Hai, motte-imasu.             [Yes, possess-STATIVE.]

5a. Sude-ni touroku-ryou o shiharaimashita ka? [Already fee OBJ-V pay-PAST QUESTION?]
5b. Shiharaimashita.              [Pay-PAST.]

6a. Kaigi ni ikimasu ka?          [Conference to  go/will-go QUESTION?]
6b. Ikimasen.                     [Go/will-go-NEGATIVE.]
```

Note that a literal translation of (4b), *"Yes, I have." is incorrect, even though it is identical to (2b), because the intended meaning is possessive, not perfective. Also note that a similar literal translation of (1b), *"Hai, watashi wa shimasu." is equally unacceptable in Japanese. In general, short answers should not be literally translated.

2 Alternatives to short answers

In English, it is possible to answer simple yes/no questions in a number of different ways:
Question: Do you have the form?

1. Minimal answer: "Yes."
2. Brief answer: "I do."
3. Short answer: "Yes, I do."
4. Long answer: "I have the form."
5. Long answer: "Yes, I have the form."

The minimal answer (1) is in common use in the armed forces, and in parts of New England. However, most other English speakers feel that it is too abrupt. Sometimes it can be used to mean that the speaker does not want to cooperate with the questioner. In certain contexts, the minimal answer can be considered rude.

The brief answer (2) does not answer the *yes/no question* directly, but rather states the information needed to determine the answer. It has formal overtones, and is used in ceremonies, when talking with a police officer, etc. It too might be considered brusque or rude.

The long answers (4 and 5) again state the information explicitly. They are much longer than required, and thus violate Grice's maxims of quantity and manner [Lev83]. For this reason, they must convey more information than the answer alone. Otherwise, the speaker sounds stilted and unnatural, as if he/she is "talking like a robot". The long answers are used in critical situations, when it is essential not to be misunderstood (child talking to angry parents, negotiator talking to hijackers, etc.) and can convey that feeling. Depending upon the tone of voice, they can also be used to express strong emotion, including e.g., regret, outrage, annoyance, or triumph. Naturally, this form can also be considered rude in everyday conversation.

The English short answer (3), although redundant, is the preferred form. It is neither too brief nor too long. It answers the yes/no question and provides an informing statement; it is long enough to suggest that the speaker is still interested in cooperating (unless, of course, extra prosody to the contrary is used).

Thus, short answers are significant because they are not rude, do not convey extra unwanted nuances, and are used in everyday conversation. It is important for a machine translation system to recognize and appropriately use short answers if it is to sound natural.[1] Note significantly that literal translations of the answers, e.g. generated in the long answer form, are completely semantically well-formed and pass most other pragmatic tests. It is only the fact that the utterance is being used in the context of question-answering that makes a normal translation unacceptable. Thus, context-sensitive processing is required.

3 The System

Output from a semantic parser is sent to the NP (Natural language Plan inference) system. The output is in feature structure format, with cyclic co-references permitted. Utterances from a dialog are processed one by one. Currently, as the parser is not yet completely finished, the actual input to the system consists of an "expected output" corpus generated by the parsing group. The NP system recognizes the short answers, marks them as such, fills in the verb, and sends the resulting structure to the transfer module. The transfer module then transfers the short answer utterance into the target language. The resulting semantic feature structure would then be sent to a generation module for text generation.

The NP system works with a set of (feature structure) assertions, and a number of plans. The assertions represent utterances in the conversation, situations, the assumed knowledge and intentions of the speakers, world knowledge, etc. The system is not restricted to processing just one input feature structure at a time, as are most rewriting systems. Thus, the system is closer in flavor to an "expert system" inference engine than to a parser. This allows context situations [BP83] to be specified and reasoned with—the system is not restricted to being context-independent.

[1] This analysis ignores the answers to more complex yes/no questions, such as conducive questions, indirect speech acts, tag questions, rhetorical questions, and focusing questions. For analyses of the questions themselves, see Kiefer [Kie80] and Quirk et al. [QGLS85].

Recognition of the occurrence of a (negative) short answer in Japanese is performed by the following two plans:

```
  [[action [[RELN short-answer-set]]]
   [dec1   [[RELN short-answer-negative-stative-set-1]]]]

[[action [[RELN short-answer-negative-stative-set-1]]]

 [prec1 [[RELN  QUESTION-IF]
         [AGEN ?questioner]
         [RECP ?answerer]
         [OBJE [[RELN  ?verb]
                [AGEN  ?agent]
                [OBJE ?REST1 ]]]]]

   [dec1 [[RELN  Iie-NEGATIVE]
          [AGEN ?answerer]
          [RECP ?questioner]]]

   [dec2 [[RELN  NEGATE]
          [OBJE [[RELN  TeIru-STATIVE]
                 [AGEN @X01[]]
                 [OBJE [[RELN  ?verb]
                        [AGEN @X01]
                        [OBJE []]]]]]]]            ;Null object.

   [eff1 [[RELN QUEUE-FOR-TRANSFER]
          [OBJE [[RELN  SHORT-ANSWER]
                 [OBJE [[RELN  NEGATE]
                        [OBJE [[RELN  TeIru-STATIVE]
                               [AGEN ?agent]
                               [OBJE [[RELN  ?verb]
                                      [AGEN ?agent]
                                      [OBJE ?REST1 ]]]]]]]]]]]
```

Plan actions have preconditions, decompositions, and effects. If all of the preconditions and decompositions are believed true, the action is recognized and the effects are asserted.

In this case, the precondition is that a yes/no question with a particular verb has been asked. (The ?tokens indicate variables, and the @tokens indicate co-reference tags.) The two decompositions specify that the person answering must respond with an "iie"["no"], and then a negative-stative statement using the same verb but no agent and no object. If all of these conditions are met, then the plan is recognized.

A plan schema can also have optional "sufficiency sets", which allow specific combinations of assertions to entail the action. Since, in Japanese, the "iie" is optional, it is possible to recognize a short answer from only a question and an appropriate verbal response. Thus, the single sufficiency set for this plan is {prec1, dec2}.

Once the plan is recognized, its effects are asserted. In this case, the effect consists of a queue-for-transfer assertion that packages the utterance as a short answer and fills in the

verb's agent from the question. (It is assumed that either the agent has been explicitly mentioned in the question, or that the reference has already been disambiguated by one of several proposed methods (e.g. [Doh89]). Proper resolution of zero pronouns is beyond the scope of this paper.)

Plans can also be hierarchical. For instance, in this example all of the different short answer types are disjunctively classified under "short-answer-set". A processing demon is attached to this plan schema. When the "short-answer-set" action is recognized, the demon fires. In this case, the demon searches for new "queue-for-transfer" assertions, prints them out, and sends them to be transferred. The demon then resets itself.

The transfer process is done using a non-monotonic rewriting system [Has89]. Nouns, verbs, tenses and aspects, etc., are transferred as usual. However, since the transfer module is informed specifically that this is a short answer, as opposed to a normal statement, it can make the verbal auxiliary explicit (if not already there), elide the verb and the objects of the verb, and send the results to generation.

The following example illustrates this process. Figure 2 shows input excerpted from the middle of a sequentially processed dialog. Figure 3 shows the output from the transfer module.

4 Conclusion

Proper transfer of short answers to yes/no questions in dialogs is important for good machine translation. This paper has described a method for recognizing and transferring short answers, and its implementation using NP, a context-dependent feature-structure-based plan inference system.

5 Acknowledgements

Hitoshi Iida managed this project and suggested using feature structures in NP. Masaaki Nagata specified the semantic feature structures used for input and output. Toshirou Hasegawa wrote the RWS rewriting system used by NP.

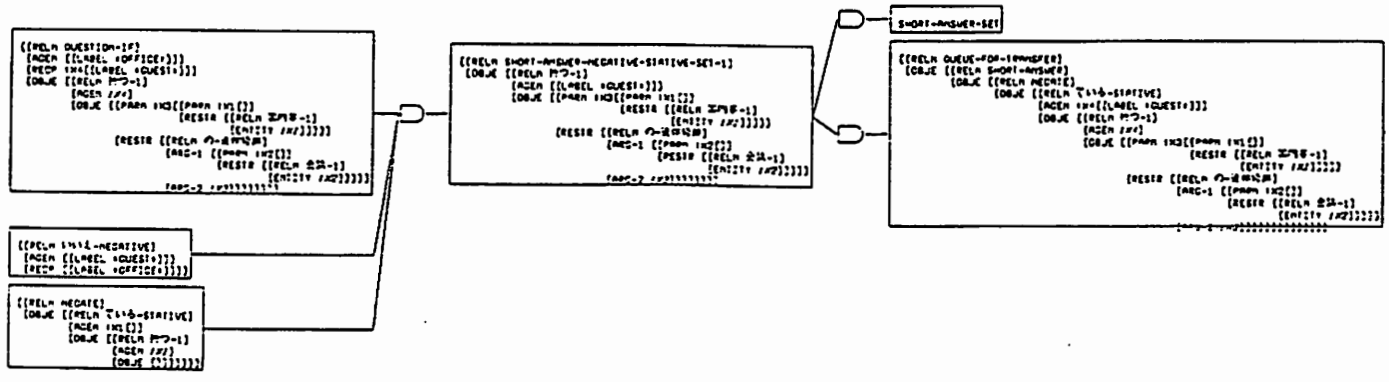

Figure 1: NP representation of the example using ATMS nodes and implications

```
"Kaigi no annaisho wa o-mochi desu ka?"
[[RELN  QUESTION-IF]
 [AGEN ?X04[[LABEL  *OFFICE*]]]
 [RECP ?X05[[LABEL  *GUEST*]]]
 [OBJE [[RELN  Motsu-1]
        [AGEN [?X05]]          ;From zero-pro. resolution
              [OBJE [[PARM ?X03[[PARM ?X01[]]
                               [RESTR [[RELN  Annaisho-1]
                                       [ENTITY ?X01]]]]]]
                    [RESTR [[RELN  No-noun-noun-1]
                            [ARG-1 [[PARM ?X02[]]
                                    [RESTR [[RELN  Kaigi-1]
                                            [ENTITY ?X02]]]]]
                            [ARG-2 ?X03]]]]]]]]

"Iie."
[[RELN  Iie-NEGATIVE]
 [AGEN [[LABEL  *GUEST*]]]
 [RECP [[LABEL  *OFFICE*]]]]

"Motte-imasen."
[[RELN NEGATE]
 [OBJE [[RELN  TeIru-STATIVE]
        [AGEN ?X01[]]
        [OBJE [[RELN  Motsu-1]
               [AGEN ?X01]
               [OBJE []]]]]]]
```

Figure 2: The Representation of a Japanese Short-Answer Exchange Used for Input

```
"Do you have the announcement of the conference?"
[[RELN QUESTION-IF]
 [AGEN [[LABEL *OFFICE*]]]
 [RECP !X4[[LABEL *GUEST*]]]
          [OBJE [[RELN DO-1]
                 [AGEN !X4]
                 [OBJE [[RELN HAVE-POSSESSIVE-1]
                        [AGEN !X4]
                        [OBJE [[PARM !X3[[PARM !X2[]]
                                         [RESTR [[RELN ANNOUNCEMENT-1]
                                         [ENTITY !X2]]]]]
                              [RESTR [[RELN OF-1]
                                      [ARG-1 [[PARM !X1[]]
                                              [RESTR [[RELN CONFERENCE-1]
                                                      [ENTITY !X1]]]]]
                                      [ARG-2 !X3]]]]]]]]]]

"No."
[[RELN NO-NEGATIVE]
 [AGEN [[LABEL *GUEST*]]]
       [RECP [[LABEL *OFFICE*]]]]

"I don't."
[[RELN NEGATE]
 [OBJE [[RELN DO-1]
        [AGEN [[LABEL *GUEST*]]]]]]
```

Figure 3: The Transferred Resulting Output Short-Answer Exchange in English

References

[AIS9] Hidekazu Arita and Hitoshi Iida. *Tri-Layered Plan Recognition Model for Dialogue Machine Translation.* TR-1-0067, ATR Interpreting Telephony Research Laboratories, Kyoto, Japan, 1989. (in Japanese).

[BPS3] Jon Barwise and John Perry. *Situations and Attitudes.* The MIT Press, Cambridge, Mass., 1983.

[dK86] Johan de Kleer. An assumption-based tms. *Artificial Intelligence*, 28(2):127–162, March 1986.

[Doh89] Kohji Dohsaka. *Identifying Zero-Pronouns Referring to Persons in Japanese Dialogue.* Technical Report TR-1-0117, ATR Interpreting Telephony Research Laboratories, Kyoto, Japan, 1989.

[Has89] Toshiro Hasegawa. *The Feature Structure Rewriting System Manual.* Technical Report TR-1-0093, ATR Interpreting Telephony Research Laboratories, Kyoto, Japan, 1989. (in Japanese).

[Kie80] Ferenc Kiefer. Yes-no questions as wh-questions. In J.R. Searle, F. Kiefer, and M. Bierwisch, editors, *Speech Act Theory and Pragmatics*, pages 97–119, D. Reidel Publishing Co., Boston, MA, 1980.

[Kno88] Craig A. Knoblock. Data-driven plan recognition. March 1988. CS Dept., Carnegie-Mellon University, Pittsburgh, PA.

[Lev83] Stephen C. Levinson. *Pragmatics.* Cambridge University Press, Cambridge, Mass., 1983.

[Mye89] John K. Myers. An assumption-based plan inference system for conversation understanding. In *WGNL Meeting of the IPSJ*, pages 73–80, Okinawa, Japan, June 1989.

[QGLS85] Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. *A Comprehensive Grammar of the English Language.* Longman Inc., New York, NY, 1985.