Dialogue Translation vs. Text Translation

-Interpretation Based Approach-

Jun-ichi Tsujii, Makoto Nagao
Dept. of Electrical Engineering,
Kyoto University,
Yoshida-honmachi, Sakyo, Kyoto, 606, JAPAN

Abstract

The authors discuss the differences of environments where dialogue translation and textual translation systems might be used. The differences are summarized as *clear definition of information* and *active participations of speakers and hearers* in dialogue translation. A new approach to MT, *interpretation based approach*, is proposed to take the advantages of dialogue translation environments. The approach introduces *a layer of understanding* to MT and can produce less structure bound translations than conventional approaches.

1 Introduction

Although we had been engaged in developing an MT system of texts for several years (Mu project [Nagao85, Nagao86]), we were puzzled when we examined the data of dialogue translation gathered by the research group of ATR, which is a newly established research organization for translation of telephone dialogues and is now gathering dialogue translation data in various hypothetical situations.

The sample translations gathered by the ATR research group looked very difficult for machines, but we rarely found syntactic structures which make textual translation so difficult, such as long noun phrases or clauses, complicated conjuncted phrases, etc.([Tsujii84] [Tsujii88]). On the other hand, most of the translations of dialogues between Japanese and English, which were produced by professional human interpreters, did not preserve syntactic structures of their original sentences at all. They were completely paraphrased in the target language and seemed very hard to be produced by conventional techniques developed for textual MT systems.

Both translations, dialogue and textual translations, are difficult, but their difficulties are very different from each other.

We discuss in this paper the differences of dialogue translation systems and textual translation systems. Bcause we do not know the difficulties of recognition of spoken utterances, we will avoid the discussion about the difficulties of interfacing the speech recognition part and the linguistic processing part, which we will certainly encounter in spoken dialogue translation systems. The dialogue translation in this paper is restricted to the translation of dialogues through keyboards, on which ATR is now concentrated.

The differences of these two translation systems mainly come from the fact that dialogues of certain types are more goal-oriented than ordinary texts. We will argue that the goal orientedness of dialogues makes dialogue translation systems more feasible than textual translation systems, though they are usually considered much harder.

2 Differences of Environments

In the current states of the art in machine translation, most researchers may agree that we cannot expect an ideal FAMT system which can translate any linguistic materials in any subject domains. So, at present, what should be discussed about MT systems have to be *engineering problems*.

We should discuss problems from engineering points of view. That is, we should discuss, first of all, what types of systems or system organizations are economically and technically feasible in what situations of actual translation, and what sorts of human aids can be expected in real application environments.

The important consideration is how to design feasible MT systems which can be used in actual, rather specific, translation *environments*. Different application environments require different technologies. Therefore, the questions we would like to pose in this paper are:

- Which is more feasible in *actual application environments*, dialogue translation systems or textual translation systems ?

- Can we design a feasible dialogue translation system just by extending or modifying current MT technologies developed exclusively for textual translation ?

Our answer to the first question, though it might sound strange, is that dialogue translation systems of certain types are more feasible than textual translation systems which are currently developed and commercially available.

It might be the case that we imagin dialogue translation is easier, because we have been engaged in developing a textual translation system and have recognized many, not only difficult but also nasty and dirty problems in textual translation systems ([Nakamura86]).

But not only because of that, we believe dialogue translation systems are more feasible, mainly because of the basic differences of environments where these two types of systems will be used.

We can summarize the differences of environments in which these two types of systems might be used as follows.

- *Clear Definition of Information* : In certain types of dialogue translations, we can define rather clearly what *information* should be transmitted from source sentences to target translations, while we generally cannot in textual translation.

  By certain types of dialogues, we mean here the dialogues such as dialogues for hotel reservation and conference registration which are currently picked up by the ATR research group, dialogues between patients and doctors tried by the CMU group ([Tomita86]),

  etc.

- *Active Participations of Speakers and Hearers* : In most application environments of textual translation systems, they are supposed to be used by *professional translators*. We cannot have the writers of texts at the time of translation, the persons who prepare texts and really want to communicate something through the texts. The actual readers of translated texts are not available, either, at the time of the translation, who really want to get messages or *information* encoded in the texts.

  On the contrary, in dialogue translation, we have both the speakers (the senders of messages) and the hearers (the receivers of messages) at the time of translating messages.

These two differences make, we claim, dialogue translation systems more feasible in actual translation environments, if they are properly designed for taking these advantages.

Our answer to the second question is directly derived from the above discussion. That is, in order to take the advantages of dialogue translation, the system organizations should be different from those for textual translation. Mere extension of current MT technologies for textual translation will not result in high quality dialogue translation systems by which *ordinary people* can communicate with each other.

We will discuss what implications the basic differences of enviroments have in the design of dialogue translation systems and, substantiate the conclusion that *if they are properly designed, certain types of dialogue translation systems, are more feasible*, technically at least, than the text translation systems which are currently available.

3   What should be translated ?

Fig. 1 shows a simplified framework of application systems of natural language understanding (NLU) other than MT systems. In this framework, *understanding of a sentence* is regarded as a process of transformation from an input sentence, a linear sequence of words, into so-called *the meaning representation of the sentence*.
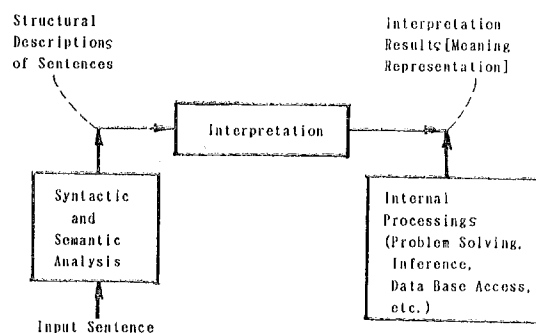

Fig. 1   General Organization of
NLU Application Systems

Meaning representation in this framework is the input to certain *internal processings* such as deductive inferences, problem solvings in certain restricted domains, data base accesses, etc., which are actually implemented as computer programs to carry out certain specific internal tasks.

Meaning of input sentences are defined in this framework, relative to the internal tasks that the systems are

expected to perform. In other words, what kinds of information should be extracted from sentences are predefined, depending on the aims of the internal processings of the systems. *Understanding* is regarded as an extraction process of information relevant to specific internal tasks.

However, the internal task or the aim of translation is

> *to re-express by using sentences of target languages the information of all aspects conveyed by sentences of source languages, with as least distortion as possible*([Tsujii86]).

The internal task of MT, by itself, does not define what information should be extracted from input texts. It is commonly recognized by linguists that all different surface linguistic expressions convey different *meaning*. MT systems, at least textual translation systems, have to extract all the factors relevant to the determination of surface linguistic expressions.

Most of the difficulties peculiar to MT, such as the selection of appropriate target lexical items or expressions, etc. come from the fact that we cannot define in MT what aspects of information in source sentences are relevant to the determination of target expressions and should be extracted from source sentences. In general, we cannot establish a representational framework which is language universal and by which *understanding* results are represented.

As a consequence, most of the current systems use certain linguistic levels of structural descriptions of source sentences, such as *deep case structures* in the Mu project, in order to calculate appropriate target descriptions. Because the structures are far from representing *understanding results* and reflect the linguistic strutures of source sentences, their translation results are inherently *structure bound*.

On the other hands, in certain types of dialogues, we can define by the purpose of dialogues what is essential or important information conveyed by uttrances and should be transmitted to their translations. Here, we do not discuss about the systems which are capable of translating arbitrary dialogues like chatterings among house wives without any purposes, but the systems which translate dialogues of certain restricted domains as already mentioned, such as dialogues for hotel reservation, conference registration, etc. In such dialogues, we can define *important information* by referring to the aim of the dialogues.

Such *important information* should be extracted from

the input and properly transmitted to the target. So, the framework for dialogue translation becomes similar to that of the other applications of NLU illustrated in Fig. 1. We can introduce a layer of explicit understanding to MT systems, to which *important information* of untterances are related and so, in which results of *understanding* can be represented in a language independent (but task dependent) way ([Tsujii87]).

Some parts of utterances which convey *information important* for the purposes of the dialogues are related to this layer and *interpreted*. Because information is expressed language-independently in this layer, we can expect *less structure bound translation results* for the parts of utterances. On the other hand, the other parts which do not convey important information need not be related to this explicit understanding layer. They would be translated by conventional MT technologies.

Let me show you a simple example from hotel reservation dialogues, which actually appears in the experiments conducted at ATR.

> [ex 1]
>
> [Japanese] hoteru(*hotel*)-wa, tomodachi(*friends*)-to Disuko(*discotheque*)-ni ikitai(*to want to go*)-node, Roppongi(*Roppongi - the name of the place in Tokyo*)-no chikaku(*to be near*)-ga iino(*to be good*)-desuga ?
>
> [Structure Bound English Translation] As for hotel, because [I] would like to go to Discotheque with friends, to be near to Roppongi is good.
>
> [English Translation] Because I would like to go to discotheque with friends, I prefer to stay at a hotel near to Roppongi.

In this example, we can divide the utterance into two. One is the part which contain important information for hotel reservation, and the other is the part which does not. Because the location of the hotel which the client wants to stay is important for the task of hotel reservation, the underlined part of the uttrance is important and should be translated as properly as possible.

The other part of the utterance, which gives the reason why the client wants to stay at a hotel in a specific region of Tokyo (Roppongi), is less important. Our contention is that these two parts of the utterance should be treated differently in dialogue translation systems.

Note that the English translation given above has a *deep case structure* completely different from that of the

source sentence. The translation contains the verbs *to prefer* and *to stay* whose corresponding Japanese verbs do not appear in the source sentence.

4 Architecture of Dialogue Translation Systems

Fig.2 shows a schematic view of a system which translates dialogues in a certain restricted domain. The translation system knows in advance what kinds of *information* or *concepts* are important for the natural flow of dialogues in that specific task domain, and also knows a set of surface linguistic expressions which may convey such *important information* .
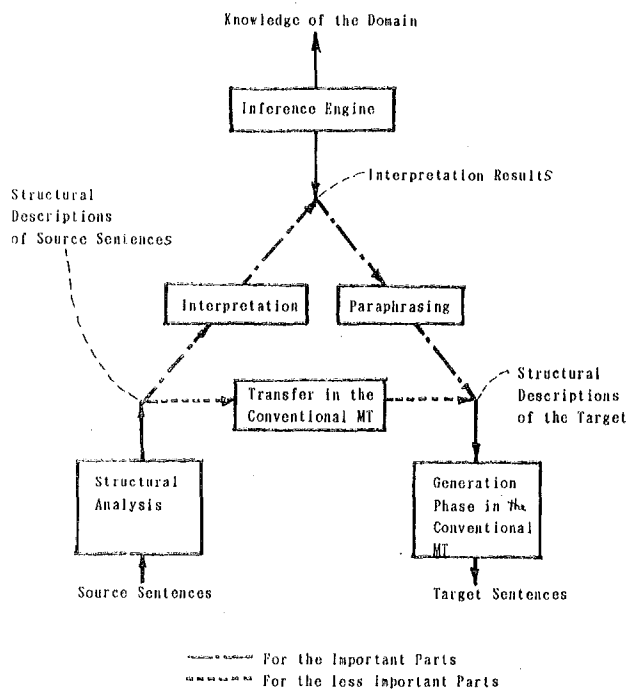
Fig. 2 The Organization of a Dialogue Translation System

By using these kinds of knowledge, the system should be able to distinguish the parts which convey important informational contents extract them and relate them to the representations of the explicit understanding layer.

It is certainly difficult to capture the important parts of untterances and *understand* them, but if we confine ourselves to a certain restricted task domain, it is much easier than story understandings in general, which AI researchers have been interested in.

Furthermore, it is easier than developing *intelligent*

dialogue systems which make conversations with human users in restricted task domains, for example, to make appropriate hotel reservation. Although those intelligent systems should be able to understand fully the user's utterances, a dialogue *translation* system needs not. The hearer, the receiver of the translated messages may understand the speaker's intention. A translation system is only required to provide *information* sufficient for his understanding. It is desirable but not inevitable for a dialogue tarnslation system to have the ability of recognizing the speaker's plan.

A translation system which extracts *important information* from source utterances and re-expresses it in the target language can produce less structure bound translations. It can reduce varieties of surface expressions to a single *meaning representation*, if they convey essentially the same information, the *same* from the view point of the purposes of dialogue. For example, the following Japanese expressions, which have quite different (deep case) structures, may be reduced to a single representation and re-expressed by English expressions. The English expressions will be chosen independently of Japanese structures but only by considering English contexts where the expressions are located.

[EX 2]

[Japanese] Roppngi-no chikaku (*to be near*) - no hoteru (*hotel*) -ga ii (*to be good*) wa.

⟶ [Structure bound translation] A hotel near to Roppongi is good.

[Japanese] Ropponi-atari (*around*) -no hoteru (*hotel*) -wo onegaishimasu (*please*).

⟶ [Structure bound translation] A hotel around Roppongi, please.

[Japanese] hoteru (*hotel*) -ha roppongi-no chikaku (*to be near*) -ga iinodesuga (*to be good*)

⟶ [Structure bound translation] As for hotel, to be near to Roppongi is good.

[Japanese] tsugou-ga-iino (*to be convenient*) - ha roppongi-ni chikai (*to be near*) hoteru (*hotel*) desuga.

⟶ [Structure bound translation] What is convenient is a hotel near to Roppongi.

As an extreme, we can imagine a system which produces fluent translations only for important parts of utterances but awkward ones for the other parts.

691

[EX 3]

[Because (to Go Discotheque) Friends] I prefer
to stay at a hotel near to Roppongi.

Note that a dialogue translation system needs not understand utterances completely, and so, it needs not understand why the clause 'tomodachi-to disuko-ni ikitai'(I would like to go to discotheque with friends) can be the reason for staying at a hotel near to Roppongi. To understand this, a system has to have a lot of real world knowledge which is not so closely related with hotel reservation tasks, such as

(1)Roppongi is a special region in Tokyo where
many discothetques exist
(2)In order to go to some place, it is preferable
to stay at a hotel near to the place
(3)If something is preferable, the client tend
to .....
etc.

A system which converses intelligently with human to make hotel reservation should have such knowledge and abilities of using it. However, a dialogue translation system has only to provide information to the human participants who organize conversation intelligently.

5 Active Participation of Speakers and Hearers

What should be *understood* from texts is highly dependent on the intentions of actual writers and readers of texts, but neither of them is available at the time of translation in textual translation.

The same texts would be read by different readers with different intentions who would like to get different sorts of information from the translated texts.

Readers of translated texts are often irritated because they cannot get necessary information for them. We found that translated texts are irritating not only because translations are awkward, but also because original texts themselves do not contain *information* which actual readers would like to get. Furthermore, evaluating translations produced by MT systems is difficult, because the evaluation highly depends on both what readers want to know and what source texts really contain. MT systems cannot produce good translations from bad source texts.

However, the *environments* of dialogue translation systems, in which both actual writers and readers are available at the time of translation, are much better than textual translation. The readers can ask questions directly

to the writers in order to get necessary information, when they cannot get it from the translated messages or when they cannot understand the translations.

Furthermore, the translation system can also pose questions to the writers (senders of messages) to clarify their intentions. We can expect an intelligent translation system to play a role of a coordinator of conversations by keeping track of exchanges of *important information* between dialogue participants (see Fig.3).

[EX 4]

[English participant] In which region do you
want to stay in Tokyo ?

[Japanese participant] disuko-ni ikitai. (I would
like to go to discotheque)

[System's Question to the participant] shitsumon-
ha anata-no kibou-suru hoteru-no basho desu
? (The question is 'in which resgion do you
want to stay in Tokyo ?'. Would you specify
the place which you prefer to stay ?)

[Japanese participant] disuko-ni chikai hoteru-
desu. (A hotel near to a discotheque)

[Translated reply to the English participant] I
prefer to stay a hotel near to a discotheque.

Note also that what is important in dialogue translation is the exchange of information through translation but not translated texts obtained as the result. Translations are satisfactory when the participants achieve their goals, even if they are awkward. On the contrary, in textual translation, translated texts themselves are important and they should be natural and clear enough in all aspects, because different readers with different intention will read them and be interested in different aspects of informtional contents of same texts.

6 Conclusions

We discussed in this paper the differences of dialogue translation systems and textual translation systems. Especially, we emphasized the differences of environments where these two types of systems will be used, and discussed what implications the differences have in the design of feasible dialogue translation systems. The main differences are:

• Clear Definition of Information in Dialogue Translation

• Active Participations of Speakers and Hearers in Dialogue Translation

We argued that, if they are properly designed to take these advantages of dialogue translation, dialogue translation systems can be more feasible than textual translation systems. Especially, we proposed a new approach to MT, called *interpretation based approach*, in which an explicit layer of *understanding* is introduced and parts of utterances conveying *important information* are *interpreted* by being related to this layer.

Though the approach produces *less structure bound* translations through *understanding and paraphrasing*, it is different from the conventional pivot or interlingual approach which claims their understanding results can be represented in the forms which are independent on both individual languages and tasks. The understanding layer in the proposed approach, on the other hand, is *language universal* but highly dependent on specific tasks of dialogues. In the proposed approach, we have to design an internal meaning representation specific to the domain of the dialogues.

The followings are important in order to develop a feasible dialogue translation system based on the interpretation based approach.

- Integration of different layers of descriptions: We have to devise technologies for integrating the descriptions of the understanding layer and the conventional structural descriptions of source sentences to produce translations, because single utterances generally consist of the parts which convey *important information* and those which do not. The idea of *safty net* should be re-considered in this new context.

- Flexible interaction during translation: Traditional post- and pre-editings by human are not the best way to take the advantage of the availability of speakers and hearers in dialogue translation. We have to design much flexible interaction modes including *clarification dialogues* between the system and the dialogue participants.

- Management of dialogue structures: In order to find *important information*, a system should have the ability of managing the dialogue sturutures. It should be able to utilize various kinds of knowledge such as knowledge about surface clue expressions, task depeadent knowledge, discourse structures, etc. to recover the structures of on-going dialogue. Especially, a translation system as a coordinator of conversations has to keep track of *important informa-*

tion exchanges through sequences of utterances.

Acknowledgement

References

[Nagao85] Nagao, M., Tsujii, J., Nakamura, J. : The Japanese Government Project for Machine Translation, Computational Linguistics, Vol. 11, NO.2-3, 1985.

[Nagao86] Nagao, M., Tsujii, J., Nakamura, J. : The Transfer Phase of the Mu Machine Translation System, Proc. of Coling 86, 1986.

[Nakamura86] Nakamura, J., Tsujii, J., Nagao, M. : Solutions for Problems of MT Parser, Proc. of Coling 86, 1986.

[Tomita86] Tomita, M., Carbonell, J. : Another Stride Towards Knowledge-based Machine TRanslation, Proc. of Coling 86, 1986.

[Tsujii84] Tsujii, J., Nakamura, J., Nagao, M.: Analysis Grammar of Japanese in the Mu-project, Proc. of Coling 86, 1986.

[Tsujii86] Tsujii, J. : Future Directions of Machine Translation, Proc. of Coling 86, 1986.

[Tsujii87] Tsujii, J. : What is PIVOT ? Proc. of MT Summit, Hakone, Japan, 1987.