Dependence on Context
in case of English-Japanese Machine Translation I - 1

Katsuyuki Shibata

Abstract

In case of a machine translation from English to Japanese, the adequate Japanese words for English pronuns "that", "this" and "you", as well as for verb "be" and interjection "hi" depend on the context inherited from the preceding sentences. Our machine translation system, US system, checks the preceding English sentence and the sentence before the last, their Japanese translations, their subject and/or object nouns, their verbs, as well as their attributes to choose the most adequate Japanese candidate for an English word in question. All the example sentences are taken from Lesson 1 of English textbooks for the first year grade Japanese junior high school students.

Key words: machine translation, context dependence, context inherited from the preceding sentences, US system.

1. Introduction.

Machine translation sysytems sold on the market carry out translations sentence by sentence and they do not analyse the context reflecting the preceding sentences. This is probably because the commercial machine translation systems presume scientific or technical reports and the like as the documents to be translated and the sentences in those articles are generally long, complicated and self-contained in themselves, and can be translated rather independently of the preceding sentences.

On the other hand, Ueno-Shibata system (US system, for brevity), the machine translation system we have been developing, simulates the man's second language aquisition process and therefore applies sentences in English textbooks used in junior high schools in Japan to the tuning up. The English sentences in such textbooks are generally short, conversational and highly dependent on the context. For example, see the question and the answer in the following conversation.

Kenji: Do you play the guitar?

Kenji: Anata wa gitaa wo hiki-masu ka?

Yumi: Yes, I do.

Yumi: hai. Hiki-masu.

Department of Applied Mathematics, Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan

None of the commercial translation systems can translate the answer sentence correctly because in Japanese, the answer sentence must repeat the verb in the question sentence. In contrast, our US system refers to the preceding question sentence and repeats the verb "hiki-masu" (play) in palce of the translation of the auxiliary verb "do". This kind of examination of the preceding sentences and the usage of the context information to generate the translated sentences are also indespensable in case of translations from English to other European languages most of which have no auxiliary verb corresponding to "do" as mentioned above.

The present article is the first one of a series of papers concerning the context dependency in case of English-Japanese machine translations. This time we explain the context analysis arising in the translation of English sentences of Lesson 1 of five text books, namely, New Horizon 1, New Prince 1, Sunshine 1, New Crown1 and New Crown1 (New Edition) for the first grade junior high school students.

2. Should "that" be translated into "are" or "sore"?

English pronoun "that" can be translated either as "are" or "sore" in Japanese.
This is a book.
That is a map. ....(1)
In case "That is ..." follows "This is ..." as above, such "that" naturally corresponds to "are" in Japanese. The case when "that" is translated into "sore" in Japanese is, for example, as follows;

> ___That___ is because computers are newer than many other things. .... (2)
> それ
>
> ___Sore___ wa konpyuutaa ga ooku-no hoka-no mono yori atarashii kara desu.

"That" above stands for what the preceding sentence asserts.

In either case, to decide whether "that" should be translated as "are" or "sore", it is necessary to examine the preceding sentence.

For sentence (1) above which is cited from Lesson 1 of a first grade junior high school textbook, the choice rule for Japanese word candidates for "that"
22;* あ れ ;2;E0=that;T0◇?;T1=v;T1◇p;J-1◇1; BS=*this is; C540;
checks the preceding English sentence (BS) and confirms that it contains the character string "this is".

For the description of the choice rules and the grammar rules, see the author's books [1] and [2].

3. Should "this" be translated into "kore" or "kochira"?

English pronoun "this" has corresponding two Japanese words "kore" and "kochira." "This" is translated as "kochira" in case of, for example, a conversation on the telephone as follows;

> Kenji: Hello. ___This___ is Kenji speaking. May I speak to Nancy?
>
> こちら
>
> Kenji: ___Kochira___ wa Kenji desu. Nanshii to hanashi-tai no desu ga.
>
> Nancy: ___This___ is Nancy. .... (3)
> こちら
> Nanshii: ___Kochira___ wa Nanshii desu.

Therefore it is prudent to confirm that the preceding English sentence does not contain "May I speak to" before we exclude "kochira" from the candidates for "this."

To the case of translation

Is ***this*** a pen? .... (4)
<u>これ</u>
Kore wa pen desu ka?

, we apply the choice rule for Japanese word candidates

22;こちら;2;E0=this;T2◇hKM;BS◇*May I speak to; D540;

and exclude "kochira" after confirming that the preceding English sentence (BS) does not contain "May I speak to". When we are translating the first English sentence in a paragraph and thus there is no preceding sentence to it, the condition mentioned above is trivially satisfied.

4. Should "be" be translated into "dearu" or "iru"?

Verb "be" is used as a copula (the corresponding Japanese is "dearu") as well as "exist" (the corresponding Japanese is "iru"). "Be" is translated into "iru" (exist) in such a sentence as follows.

Mother ***is*** in the kitchen. .... (5)
<u>います</u>
Okaa-san wa daidokoro ni ***i-masu***.

In the following question-answer dialogue,

Is this a pen?

Kore wa pen desu-ka?

Yes, it ***is***. .... (6)
<u>です</u>
Hai, sou ***desu***.

"is" in answer (6) is verified not to mean "exist" by the choice rule

33;いる ;2;T0=v;E0◇be ;T0◇?;J-3◇7;TA◇ah; PJS◇*にいますか？; G635;

confirming that the preceding Japanese translation (PJS) does not contain "ni i-masu ka?"

By the way, we translate "be" verb into "aru" in Japanese for the existence of inanimate objects or of plants. Our US system treats the cases for this usage of "be" in another way.

5 . "Be" as a verb or an auxiliary verb?

"Be" can also be used as an auxiliary verb in a progressive form sentence. For example, in the following dialogue

Is he playing the gutar?
<u>います</u>
Kare wa gitaa wo hiite-***i-masu*** ka?

Yes, he ***is***. .... (7)
<u>弾いています</u>
Hai, ***hiite-i-masu***.

, the answer in Japanese repeats the verb "hiite-i-masu" (is playing) in the preceding question sentence. Accordingly, even in the case of such a dialogue as follows:

Is he a junior high school student?
<u>です</u>
Kare wa chuugakusei ***desu*** ka?

Yes, he ***is***. .... (8)
<u>そうです</u>
Hai, ***sou-desu***.

, it is necessary to verify that "is" in (8) is not an auxiliary verb as in (7) although sentences (7) and (8) are completely identical in themselves. In such a sentence as (7) in a dialogue, the variable VZ which preserves the attributes of the verb in the preceding sentence or VZ2 which preserves those of the verb in the sentence before the last contains the attribute symbol 'G' indicating that the verb in question

appeared in the progressive form (-ing form). Therefore the generation rule for the translation of the example sentence (8)

3;6023;X;0;FT;LT;T0◇N;T3=v;T3◇ijk;VZ◇G; VZ2◇G;J2#Z;J3#Z;J3#+そうです; 6350;

confirms that neither the variable VZ nor VZ2 contains the attribute symbol 'G' before adding "sou-desu" to the Japanese translation.

6．Is "you" singular or plural?

Concerning "you" which appears in a conversation, it is often very difficult to know whether that "you" means singular or plural. This problem is explained in detail in the author's book [1], p.84-85. In order to resolve the ambiguity as correctly as possible, we proposed in [1] to check verious factors, one of which is the preceding English sentence. Looking up the examples of plural "you" appearing in junior high school English textbooks, we find:

Hachi: ***Tell us***. Come on, Zen-san.
　　　私達に教えなさい
Hachi:　***Watashitachi ni oshie-nasai***. Saa, Zen-san.

Zen: I'll tell ***you*** if you promise that ***you*** won't laugh. .... (9)
　　　あなた達
　　　　　　　　あなた達

Zen: ***Anatatachi*** ga warawa-nai darou to iu koto wo yakusoku-suru naraba watashi wa ***anatatachi*** ni oshie-mashou.

, or

Now you ***tell us*** "What did they say?"

　　　私達に言いなさい
Saa, anata wa "Karera wa nan-to itte-i-mashita ka" wo ***watashitachi ni ii-nasai***.

You have half a minute to write your answers. .... (10)
あなた達　　　　　　　　　　30

Anatatachi wa kotae wo kaku tame-ni 30 byou ga ari-masu.

---

and the like. In some cases we see "tell us" in the sentence preceding the one containing "you" (plural). So we decided to confirm that the preceding sentence does not contain "tell us" before excluding the plural interpretation (anatatachi) from the Japanese candidates for "you". The following is an example of such a case:

Bert: Are ***you*** from Japan, Taro?
.... (11)
　　　あなた

Baato: ***Anata*** wa Nihon kara ki-mashita ka, Tarou?

In the choice rule applied to "you" in (11)

22; あなた達 ;2;E0=you;E-3◇everyone;E-7◇ teacher;BS◇*tell us;TA◇g; D880;

we find the condition that the preceding English sentence (BS) does not contain "tell us". Of course, if "you" in (11) were certain to correspond to "Taro", since "Taro" is a person's name, the plural interpretation for "you" would be excluded. But in cases like (11), 'you' may well correspond to a group of persons including Taro, and it is not easy to obtain a definite conclusion.

7. Japanese expressions for "hi"

We have three Japanese expressions "yaa", "kon-nichi-wa" and "okaeri-nasai" for English interjection "hi" registered in our system dictionary. It is indispensable to analyse the context in order to determine which one to select in

a given sentence. The following example is a case where "okaeri-nasai" should be chosen:

> Mrs. Reed: Ellen, Dad is home!
>
> Riido Fujin: Eren, Papa ga kaette-ki-mashita yo.
>
> Ellen: **_Hi_**, Dad. .... (12)
> 　　　お帰りなさい
> Eren: **_Okaeri-nasai_**, Papa.

Consequently we exclude "okaeri-nasai" from the candidates for "hi" when the preceding English sentence (BS) does not contain "is home!". This is embodied in choice rule:

XX;お帰りなさい;2;E0=hi;BS◇* is home!; e615;

8．One word sentence which doesn't repeat a word in the preceding sentences

When we translate an interrogative sentence consisting of a single noun (or a noun phrase), we distinguish from the other cases the case where the noun in question has appeared as the object word in the preceding sentence (OJ$) or in the sentence before the last (OJ2$). In the former case, we translate as

> A: Your town? .... (13)
> A 　　　　　ですか
> A: Anata-no machi **_desu ka_**?

adding "desu ka" to the noun. In the latter case, we generate the translation without adding "desu ka", simply repeating the noun in question. The generation rule applied to (13) above is

3;1;X;0;FT;LT;T0=O;T0=dghxyB;T0◇K;DJ0◇OJ$;DJ0◇OJ2$;DJ1=？;DJ0◇」;DJ-1◇、;J0#+ですか; 6117;

which confirms the the Japanese word (DJ0) for the noun in question is not the object of the preceding sentence (OJ$) nor of that of the sentence before the last (OJ2$).

9．Conclusion

As we have hitherto seen, it is necessary to analyse the context inherited from the preceding sentences by investigating the preceding English sentence and the sentence before the last, their Japanese translations, as well as the subjects, the objects and the verbs in the preceding sentences and their attributes in order to translate as correctly as possible the sentences presented in Lesson 1's in the first grade junior high school textbboks.

References

[1] Katsuyuki SHIBATA: English-Japanese Translation System by C Language - How to Construct and tune it up, Rassel Inc., T okyo, 1990.
http://www1.rsp.fukuoka-u.ac.jp/chosho/cgen_j.html

[2] Katsuyuki SHIBATA: English-Japanese Translation System / Example Sentences and Tuning Up, Rassel Inc., Tokyo, 1994.
http://www1.rsp.fukuoka-u.ac.jp/chosho/eiwa_j.html