

Dependence on Context in case of English-Japanese Machine Translation I – 13

Katsuyuki SHIBATA¹⁾

(Received November 30, 2010)

英和機械翻訳における文脈依存性 I – 13

柴田 勝征¹⁾

(平成 22 年 11 月 30 日受理)

Abstract

This is the thirteenth of a series of articles on the context dependency analysis in the case of English-Japanese machine translation system which we call “US system”. In this issue we examine the examples taken from Lesson 13 of three English textbooks for the first year grade Japanese junior high school students; namely, Sunshine 1 (Kairyudo), New Crown 1 (Sanseido) and New Horizon 1 (Tokyo Shoseki).

The main subjects to be examined in this issue are the followings: (1) Does “program” mean “bangumi” of TV or “puroguramu” for computers? (2) Replace the translation of the substitute verb “be” by that of the verb in progressive form in the preceding sentence. (3) Japanese translation of “them” determined by the subject of the preceding sentence. (4) A case where the Japanese translation of a pronoun subject is followed by the postposition “ga”. (5) Selection for the verb “play” (in case “...suru” is selected). (6) Selection of Japanese “sorera” for “those” in order to avoid repetition of a noun in plural form. And (7) does “library” mean “toshokan” or “toshoshitsu”?

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

1. Does “program” mean “bangumi” of TV or “puroguramu” for computers?

English word “program” can be translated into Japanese “bangumi” of TV/radio or “puroguramu” for computers.

I'm in a TV station. 私はテレビ局にいます。 Watashi wa terebi-kyoku ni i-masu.

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

福岡大学理学部応用数学科, 〒 814-0180 福岡市城南区七隈 8-19-1

I'm watching TV.
 私はテレビを見ています。
 Watashi wa terebi wo mite-i-masu.

We can see many *programs* here. (1)
 私達はここで多くの番組を見る事が出来ます。
 Watashitachi wa koko de ooku-no *bangumi* wo
 miru koto ga deki-masu.

In the example sentence (1) above, the Japanese translation of the word “program” is chosen among “puroguramu” and “bangumi” by the following choice rule;

11;* 番組; 2; E0=program; E-1=many; E-2=see; OJ=テレビ ;7607;

The choice rule above checks that the word is preceded by “see many” and that the object of the preceding sentence was TV, and concludes that the adequate Japanese should be “bangumi” here according to the context.

2. Replace the translation of the substitute verb “be” by that of the verb in progressive form in the preceding sentence.

In the process of generating the Japanese translation for an answer to a question using a verb in progressive form, the Japanese translation of the substitute verb “be” is replaced by that of the verb in progressive form in the preceding sentence as follows.

Are they swimming in February?
 2月に彼等は泳いでいますか?
 Ni-gatsu ni karera wa oyoide-i-masu ka?

Yes, they *are*. (2)
 はい、泳いでいます。
 Hai, *oyoide-i-masu*.

The generation rule for this Japanese translation is as follows.

3; 6023; X; 0; FT; LT; T3=v; T0<>N; VZ=G; T3<>p; J 2#Z; J3#Z; J3#+VR\$; J3#-格; J3#SINKO; J3#DES; A

075;

In the generation rule above, commands #Z and #+VR\$ successively applied to the translation of the third word annihilates the Japanese translation of the word and then adds the Japanese expression of the verb in the preceding sentence kept in the variable VR in the indefinite form. The command #SINKO makes the Japanese verb into the progressive form and #DES adds the Japanese auxiliary verb meaning politeness.

3. Japanese translation of “them” determined by the subject of the preceding sentence.

In our previous article [3], we presented an example of a choice rule for the Japanese for “them” which refers to the information about the object of the preceding sentence. The analogous choice rule for “them” which we present below uses this time the information about the object of the preceding sentence.

All these elephants are working.
 これらすべての象は働いています。
 Korera subete no zou wa hataraitte-i-masu.

Who trains *them*? (3)
 それら（それらの象）を誰が訓練しますか？
 Sorera (sorera no zou) wo dare ga kunren-shi-
 masu ka?

The choice rule for the Japanese candidates for “them” in the example sentence (3) above is this:

22; 彼; 2; E0=them; E-1<>cost; E-1<>help; SZ<>h; SZ=FxX; OZ<>FxX; DJA<>。 ; J055;

The rule checks that the attribute set SZ of the subject of the preceding sentence contains “FxX” meaning “plural” but that OZ of the object of the preceding sentence doesn't. And the rule also checks that SZ doesn't contain the “human” attribute 'h' and thus rejects the Japanese translations of “them” with “human” attribute, or equivalently “彼等” and “彼女達”.

4. A case where the Japanese translation of a pronoun subject is followed by the postposition “ga”.

In a Japanese sentence the subject is followed by either postposition “wa” or by postposition “ga”. The choice between “wa” and “ga” is a very delicate question and usually depends on the context.

Look at the example sentences below.

Who cleans the city?
誰がその都市をきれいにしますか？
Dare ga sono toshi wo kirei-ni-shi-masu ka?

Everyone does..... (4)
みんながきれいにします。
Minna ga kirei-ni-shi-masu.

The generation rule for the translation of the sentence (4) above is as follows.

3; 24; X; 0; LT; E1<>can; T1<>pN; J-1<>07; J-2<>3; BS<>*doesn't; J0#+ が; J1#Z; J1#+VR\$; J1#DES; 9652;

This rule checks that the auxiliary verb is not “can” and that it is not of the past tense nor is it in negative form. The rule further checks that the preceding English sentence does not include “doesn't”. All these conditions satisfied, the postposition “ga” is added to the Japanese for the pronoun subject and the Japanese for the auxiliary verb is replaced by that of the main verb of the preceding sentence. The last condition excludes the cases like this;

He *doesn't* run.
彼は走りません。
kare wa hashiri-masen.

He does!
彼は走ります！
Kare *wa* hashiri-masu!

5. Selection for the verb “play” (in case “...suru” is selected).

Concerning the selection of the Japanese for the verb “play”, we have seen in our previous articles in this series, a case where “hiku” is rejected was presented in [1] and a case where “puree-suru” was rejected was shown in [2]. This time we exhibit a case where the Japanese translations “...suru” are adopted.

Look.
見なさい。
Mi-nasai.

He is playing shogi now.
彼は今将棋をしています。
Kare wa ima shougi wo shite-i-masu.

He is *playing* with my grandfather.(5)
彼は私の祖父とプレーしています。
Kare wa watashi no sofuo *puree-shite*-i-masu.

A choice rule for “play(-ing)” in (5) above is as follows;

33;*する; 2; E0=play; T0=G; BS=* playing; VR=する; S912;

This rule checks that the preceding English sentence contains “playing” and that the Japanese for the main verb of the preceding sentence is “...suru”. These conditions satisfied, the Japanese candidate with “...suru” are adopted, namely “puree-suru” and “suru”.

There is one more point to be paid attention to in generating the Japanese translation of (5).

The Japanese for English preposition “with” is usually “to issho-ni” in a context like this. But in a context where the expression “play with” is used in the sense of “fight against”, “to issho-ni” should be replaced by simple “to”.

4; 71; 6; 1; E0=with; LT; T1=h; E-1=play; T-1=m; PJS=*将棋を; J0#- 一緒に : issho-ni; C315;

The generation rule above verifies that the

prepositional phrase in question, “with” plus ‘human’, is preceded by “play” and that the Japanese for the preceding sentence contains “shogi wo”. When these conditions are satisfied, it deletes “issho-ni” from the Japanese “to issho-ni” for “with”.

6. Selection of Japanese “sorera” for “those” in order to avoid repetition of a noun in plural form.

See the following example sentences.

Mike: Is that tape recorder yours, too?
 マイク 「あのテープレコーダーもあなたのものですか？」
 Maiku “Ano teepu-rekoodaa mo anata no mono desu ka?”

Ken: No. It is Midori's.
 健 「いいえ。それ（そのテープレコーダー）はみどりのです。」
 Ken “Iie. Sore (sono teepu-rekoodaa) wa Midori no desu.”

Mike: Are *those* tapes yours or Midori's?..... (6)
 マイク 「それらのテープはあなたのものですか、それともみどりのですか？」
 Maiku “*Sorera no* teepu wa anata no mono desu ka, sore-to-mo Midori no desu ka?”

Ken: *They* are Midori's. (7)
 健 「それら（それらのテープ）はみどりのです。」
 Ken “*Sorera* (sorera no teepu) wa Midori no desu.”

The pronoun “those” in (6) above has two Japanese candidates, namely “arera-no” and “sorera-no”. The following choice rule selects “sorera-no” as adequate to the context, verifying that the Japanese for the noun next to “those” (DJ1 = “teepu”) coincides with that of the subject of the sentence before the preceding sentence (SJ2 \$ = “teepu-rekoodaa”).

88;* それらの ;2; E0=those; J1=1; DJ1=SJ2 \$; v530;

In fact “teepu” (DJ1) and “teepu-rekoodaa” (SJ2\$) are different words. But our equality criterion for the

Japanese words in the description of the choice rules is defined to be satisfied if one of them is included at the head or at the tail of the other as well as the case where the two words entirely coincide. This ambiguity misled the selection for “those” in (6). If we had added a condition “SZ2=xFX” which implies that the attribute set SZ2 of the word in question (SJ2 \$) contains the attributes meaning “plural”, our translation system would not have misjudged.

In contrast, the selection of the correct Japanese for “they” in (7) was adequately carried out by the following choice rule.

22;* それら ;2; E0<>those; T0<>o; T1<>u; SZ=xXF; SZ<>h; SZ2<>xX; OZ<>xXF; JA<>0; EA<>people; K590;

The choice rule above verifies that the attribute set SZ of the subject of the preceding sentence contains “plural” attributes “xXF” but doesn't contain “human” attribute 'h'. We have previously presented a choice rule for “they” in [1] analogous to the rule below, which verifies that the attribute set OZ of the object of the preceding sentence contains “plural” attributes “xXF” but doesn't contain “human” attribute 'h'.

22;* それら ;2; E0=they; FT; T0=O; SZ<>xF; OZ=xF; OZ<>h; SZ2<>h; OZ2<>h; T1<>u; K145;

7. Does “library” mean “toshokan” or “toshoshitsu”?

We have to distinguish two Japanese words “toshokan” and “toshoshitsu” for “library”. The word “toshokan” means an entire building used as a “library” as a whole, while “toshoshitsu” is a “library room” in a school or a municipal center, etc.

Ken: Look at this picture. This is my school. My classroom is on the second floor.
 健 「この写真を見なさい。これは私の学校です。私の教室は2階にあります。」
 Ken “Kono shashin wo mi-nasai. Kore wa watashi no gakkou desu. Watashi no kyoushitsu wa ni kai ni ari-masu.”

Mike: Where is the *library*? (8)

マイク 「図書室はどこにありますか？」

Maiku “*Tosho-shitsu* wa doko ni ari-masu ka?”

Ken: It is on the same floor. It is near my classroom. We have a lot of English books

in the *library*. (9)

健 「それは同じ階にあります。それは私の教室の近くにいます。その図書室に多くの英語の本があります。」

Ken “Sore wa onaji kai ni ari-masu. Sore wa watashi no kyoushitsu no chikaku ni ari-masu. Sono *tosho-shitsu* ni ooku-no Eigo no hon ga ari-masu.”

The choice rule applied to select the Japanese for “library” in (8) is as follows.

11;*図書室;2;E0=library;BS=*classroom;PJS=*階に;5090;

The rule checks that preceding English sentence contains “classroom” and the Japanese translation of the preceding sentence contains “gai ni” (on the ...th floor). These conditions indicate that the “library” in question is a room in a school.

For the selection of the Japanese for “library” in example (9) above, another choice rule below is applied.

11;*図書室;2;E0=library;PLS;F1JT=1B;F2E=classroom;5095;

This choice rule checks that the English part which is under translation contains plural sentences and that the first sentence contains a noun which has “place” attribute 'B' and that the word is in fact “classroom”. When all these conditions are satisfied, the rule concludes that the Japanese for “library” in the second sentence in the same part should be “tosho-shitsu”.

References

- [1] Katsuyuki SHIBATA: Dependence on Context in case of English-Japanese Machine Translation I-8. Fukuoka University Science Reports, vol.38 No.2. pp.63-68, 2008
<http://www1.rsp.fukuoka-u.ac.jp/chosho/Cntxt1-8.html>
- [2] Katsuyuki SHIBATA: Dependence on Context in case of English-Japanese Machine Translation I-10. Fukuoka University Science Reports, vol.39 No.2. pp.141-148, 2009
<http://www1.rsp.fukuoka-u.ac.jp/chosho/Cntxt1-10.html>
- [3] Katsuyuki SHIBATA: Dependence on Context in case of English-Japanese Machine Translation I-11. Fukuoka University Science Reports, vol.40 No.1. pp.79-87, 2010
<http://www1.rsp.fukuoka-u.ac.jp/chosho/Cntxt1-11.html>

