Dependence on Context
in case of English-Japanese Machine Translation I- 6

Katsuyuki Shibata*

Abstract

This is the 6-th of a series of articles on English-Japanese machine translation system called US system. This time we treat Lesson 6 of the first year grade 5 textbooks from 3 publishers.

It is often said that mathematics is a subject to be learned step by step and that it is consequently impossible to continue learning math if one failed to understand its basics. It seems however that junior high English is really a subject to learn step by step. The choice rules for candidate translated words and the grammar rules for generating translated phrases used in Lesson 1 are also used in Lesson 2 and thereafter, the rules used in Lesson 1 and 2 are also used frequently in Lesson 3 and thereafter, and so on. In this paper, we only treat the phenomena which appear firstly in Lesson 6.

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

1．The Choice of the Japanese words for "they" and "their"

She is an English teacher.
彼女は英語の先生です。
Kanojo wa Eigo no sensei desu.

We like her class.
私達は彼女の授業（クラス）が好きです。
Watashi-tachi wa kanojo no jugyou (kurasu) ga suki-desu.

Are Ken and Kumi in the same class?
健と久美は同級生ですか？
Ken to Kumi wa doukyuusei desu ka?

Yes, **_they_** are. .... (1)
はい，そうです。
Hai, sou desu.

Who is **_their_** homeroom teacher? ....(2)
彼等のホームルームの先生は誰ですか？
Karera no hoomuruumu no sensei wa dare desu ka?

Miss Kato is. .... (3)
加藤（加東）先生です。
Katou **_sensei_** desu.

As explained in the preceding article [2], our US System has "karera", "kanojo-tachi",

---

* Department of Applied Mathematics, Faculty of Science, Fukuoka University, Japan
福岡大学理学部応用数学科

"hito-bito" and "sorera" as the translations for "they". We have hitherto seen the examples in which "sorera", "hito-bito" and "karera" are adopted respectively. This time we are going to see an example in which "kanojo-tachi" is adopted. The choice rule for "they" in (1) above is

22;*彼女;2;SZ=W;SZ=F;OZ<>x;E000;

and it verifies that the attribute set SZ of the subject in the preceding sentence contains 'W' (Woman) and 'F' (conjunctive "and" or "or") and that the attribute set OZ of the object in the preceding sentence doesn't contain 'x' (plural form), and consequently choose "kanojo-tachi". In this example, the subject of the preceding sentence is "Ken and Kumi" and "Kumi" has attribute 'W' while "and" has attribute 'F'. We might have adopted "karera" for a group containing both man and woman like "Ken and Kumi", but our system takes the "feminist" view point of choosing "kanojo-tachi" in cases like this. Nevertheless this "kanojo-tachi" is deleted by a grammar rule during the process of the translation, and the final result becomes "Hai, sou-desu."

As the Japanese translation of "they" varies as above, so do those of "them" and "their". In example (2) following (1) above, the choice rule applied to decide the Japanese for "their" is;

88;*u;1;E0=their;E1<>impact;SO=h;SO= FxX;d620;

This rule checks that the word following "their" is not "impact", and that the attribute set SO of the subject and the object of the preceding sentence contains 'h' (human) as well as 'F', 'x' or 'X' (each of the three implies 'plural'), and concludes

to select the candidates with attribute 'u' (human). Before the application of this choice rule, the candidate "kanojo-tachi no" is already abandoned by another rule, and consequently the translation for "their" is determined to be "karera no". It may be a formal contradiction that, while "they" in (1) and "their" in (2) both refer to same "Ken and Kumi", "they" in (1) is translated as "kanojo-tachi" and "their" in (2) as "karera no". Our system does not require the extreme rigidity whether to call "kanojo-tachi" or "karera" a human group consisting of both male and female persons. This fits the linguistic reality of human communications.

2. Example of translating the title "Mr.", "Mrs." or "Miss" as "Sensei" (=teacher)

We pointed out in the third article [1] of this series that in the "classroom English" described in junior highschool textbooks, the title "Mr.", "Mrs." or "Miss" of a teacher is translated in "Sensei" (=teacher) in Japanese. This time we exhibit an example where "Miss" is translated as "Sensei".

For sentence (3) of the preceding section, the Japanese for "Miss" is selected by the following choice rule;

88;*先生;2;T0=M;FT;J3=0;OZ=e;SJ2=先生; d700;

This rule applies to the words with attribute 'M' (=Mr., Mrs., Dr. etc) which are at the head of a sentence (FT=FirsT). When the attribute set OZ of the object of the preceding sentence contains 'e' (education or culture) and the Japanese translation SJ2 of the subject of the sentence before the last is "sensei", the rule selects Japanese "sensei" for the English

word in question. In fact, the values of SJ and SJ2 are respectively "Ken and Kumi" and "sensei".

We remark that in case where there is no subject noun (i.e. pronoun subject) in the preceding sentence or in the sentence before the last, the existing value of SJ or SJ2 remains as it is.

3．Should "there is/are" be translated as "iru" or "aru" ?

In some first grade junior high school English textbooks, sentences containing "there is/are" start to appear in Lesson 6. When there is/are (a) human being(s) or animal(s), we say "ga iru" in Japanese, while the existence of inanimate(s) or plant(s) is expessed as "ga aru". In the usual answer sentence to an inquiry sentence, the subject is not mentioned in Japanese. Therefore we are obliged to refer to the preceding question sentence. Let's have a look at an example.

B：Are there many clubs in your school?
B「あなたの学校に多くのクラブがありますか？」
B「Anata no gakkou ni ooku-no kurabu ga ari-masu ka?」

T：Yes, *there are*. Ellen, Roy, and I belong to the International Club. .... (4)
T「はい，あります。エレン，ロイ，そして私は国際クラブに入っています。」
T「Hai, *ari-masu*. Eren, Roi, soshite watashi wa kokusai-kurabu ni haitte-i-masu.」

B：What do *you* do in the International Club? .... (5)
B「国際クラブであなた達は何をしますか？」
B「Kokusai-kurabu de *anata-tachi* wa nani wo shi-masu ka?」

In the example (4) above, the choice rule for "there are" is as follows;

33;u;1;T0＝ E;FT;SZ ＜ ＞ ah;JZ ＜ ＞ 1;TZ<>ah; H060;

This choice rule applies to the words (phrases) with attribute 'E' (= there is/are/ exist/stand) which are at the head of a sentence (FT = FirsT). When the attribute set SZ of the subject of the preceding sentence doesn't contain 'a' (animal) nor 'h' (human), and if there is no noun (part of speech code '1') or a word with attribute 'a' or 'h' in the successive part of the sentence, the candidate with attribute 'u' (human) is rejected. This is equivalent to rejecting "ga iru". At this stage, the recorded informations of the preceding subjects and their attributes are;

sj＝１クラブ(club)　sz＝?xxeByつ
sj2＝２サッカー部(soccer club), sz2＝emyつ

and consequently, no human or animal noun is recorded.

4．An example where "you" is interpreted as plural

In example (5) of the preceding section, "what do you do" is a phrase which has three Japanese translations in the system dictionary, i.e. "anata wa nani wo shi-masu ka", "Anata wa nani wo shi-te-i-masu ka"

and "Anata-tachi wa nani wo shi-masu ka". The choice rule applied to (5) to select one of them is;

XX;*あなた達は;2;E0=what do you do;J-1=ǀ;BS=*, and I;SZ=h;SZ=F;e772;

This rule verifies that the preceding English sentence (BS) contains ", and I", and that the attribute set SZ of the subject of the preceding sentence contains 'h' (human) as well as 'F' (conjunction 'and' or 'or'), and consequently selects the plural form "anata-tachi".

5. Should "city" be translated as "toshi" or "machi"?

English common noun "city" is usually translated as "toshi" in Japanese, but there is also a case where "city of Edo" is translated as "Edo no machi". If the "city" in the preceding sentence is translated as "machi" according to the context, the "city" in the next sentence has the possibility of also beeing translated as "machi". Therefore in a choice rule which select "toshi" for "city", we check the condition that "the Japanese translation of the preceding sentence doesn't contain "machi". Here is an example;

> Everyone is very busy.
> みんながとても忙しいです。
> min-na ga totemo isogashii desu.
>
> It's an interesting _city_. .... (6)
> それはおもしろい_都市_です。
> Sore wa omoshiroi _toshi_ desu.

The choice rule applied to "city" in (6) above is the following.

11;街;2;E0=city;E1 < > of;E-1 < > this;EA < > part;TA < > g;PJS < > 街;0825;

This rule rejects the candidate "machi" after verifying several conditions, one of which checks that the preceding Japanese translation (PJS) should not contain "machi".

6. Should the verb "watch" be translated as "wo miru" or "no ban wo suru"?

In our system dictionary, 'watch' as a verb has two Japanese "wo miru" (look at ～) and "no ban wo suru" (stand guard over ~).

> Alice _watches_ Humpty. ……… (7)
> アリスはハンプティーを_見ます_。
> Arisu wa Hanputii wo _mi_-masu.

The choice rule applied to "watch" in (7) above is;

33;*を見る;2;E0=watch;T0 < > pGP;J1=1;T1=h;T1=K;J2=0;VZ < > ijk;O522;

If someone "watches" someone (proper noun T1 = h;T1 = K;), and if the attribute set VZ of the verb of the preceding sentence contains neither 'i' (transitive verb), 'j' (intransitive verb) nor 'k' (double transitive verb), then the rule above choose "wo miru" as the adequate translation. It is of course possible that someone may stand guard over someone, but in such a case, there should precede some sentences explaining the situation, and consequently the attribute set VZ contains 'i', 'j' or 'k'. If there is no such explanations and someone is said to 'watch' someone, the choice rule above judges that "wo miru"

has a greater possibility than "no ban wo suru".

Example sentence (7) above is cited from "Alice in the Wonderland", and it has naturally the preceding context, but our US System cancels the values of all the variables concerning the information on the preceding contexts when a new chapter or section begins. Sentence (7) is placed in such a starting position and hence there is no context information.

References

[1] Katsuyuki SHIBATA: English-Japanese Translation System by C Language - How to Construct and Tune It Up, Rassel Inc., Tokyo, 1990.
http://www1.rsp.fukuoka-u.ac.jp/chosho/cgen_j.html

[2] Katsuyuki SHIBATA: Dependence on Context in case of English-Japanese Machine Translation I-5. Fukuoka University Science Reports, vol.37 No.1, pp.93-103, 2007
http://www1.rsp.fukuoka-u.ac.jp/chosho/Cntxt1-5.html