Dependence on Context
in case of English-Japanese Machine Translation I-3

Katsuyuki Shibata

Abstract

Continuing the investigation on the role of the context analysis in machine translation from English to Japanese, we treat the following three questions; (1) Should Mr., Mrs., Miss and so on be translated as "Sensei"? (2) Is Ken a Japanese male's name or the nickname for Kenneth in case it is preceded by a proper noun? (3) Which word does "too" point out when a noun in the preceding sentence is repeated in the present sentence? Questions similar to (2) and (3) were already considered in our preceding article [3], but we look these questions from a little bit different points of view this time.

In the appendix, we give a short explanation on the context checking mechanism in NTT's Japanese-English translation system ALT/JE.

1. Introduction.

Following the line of article [3], we investigate the necessary information on the preceding and the last-but-one sentences in case of the machine translation of the English sentences, this time, in Lesson 3 of the 5 textbooks for the first grade junior high school students, New Horizon, New Prince, Sunshine, New Crown and New Crown (New Edition).

2. Should "Mr.", "Mrs.", "Miss" etc be translated as "Sensei"?

In junior high school textbooks, we often find conversations between a teacher and students in a classroom. In such a case, "Mr. XXX" or "Mrs. YYY" should be translated as "XXX Sensei" or "YYY Sensei". But it is not always necessary to call a person who is known to be a teacher with "Sensei". Only when a teacher talks with his / her pupils, the teacher is called by his name with respected "Sensei" attached.

In the example below, although Mr. Mori is a teacher as profession, he is Keiko's father and a friend's father to Nancy. Therefore he need not be called "Mori Sensei" in such a situation:

Keiko: Nancy, this is my father.
恵子（景子，慶子）「ナンシー，これは私の父です.」
Keiko「Nanshii, kore wa watashi no chichi desu.」

Mr. Mori: Hello, Nancy.
森*氏*「こんにちは，ナンシー.」
Mori **Shi**「Kon-nichi-wa, Nanshii.」

Nancy: Are you a teacher?
ナンシー「あなたは先生ですか？」

Department of Applied Mathematics, Fukuoka University, Japan

Nanshii「Anata wa sensei desu ka?」

Mr. Mori: Yes, I am. I am an English teacher.   ....(1)
森*氏*「はい，そうです．私は英語の先生です．」
Mori *Shi*「Hai, sou desu. Watashi wa Eigo no sensei desu.」

In the following choice rule for Japanese candidates for "Mr." in example (1), the fact that "Mr. Mori" appeared at the head of the last-but-one sentence (EB2), the fact that Japanese proper noun "Mori" is included in the translation of the sentence before the last (JS2) but that "Sensei" is not (i.e. the Japanese sentence before the last begins with "Mori Shi / San") are verified and then the candidate "Sensei" is abandoned.

88 ; 先生 ; 2 ; E0=Mr. ; J1=1 ; T1=K ; T1=h ; E0= EB2 ; JS2=*DJ1 ; JS2<>*先生 ; d170 ;

And see the following example,

Mrs. Mori: This is your seat, Nancy.
                    ....(2)
森*夫人*「これはあなたの席です，ナンシー．」
Mori *Fujin*「Kore wa anata no seki desu, Nanshii.」

Since neither the attribute set of the subject nor that of the object of the preceding sentence (SO) contains 'e' (=education or culture) and since there is no word with attribute 'e' in the present sentence, the default choice rule below excludes "Sensei" from the candidates for "Mrs."

88 ; 先生 ; 2 ; T0=M ; SO<>e ; TA<>e ; TZ<>e ; JZ<>% ; d810 ;

3. Is Ken a Japanese male's name or a nickname for Kenneth in case it is preceded by a proper noun?

Until recent years ago, we used to say a Japanese's first name and family name in the same order as in the case of English and Americans, but nowadays we say the family name first followed by the first name just as we say them in Japanese. (see the discussion in "Word-Quest Mail, No. 269 [4]).

The trend is reflected to Japanese junior high school textbooks, and many English textbooks write a Japanese' family name and first name in this order. The first sentence in the example below is such a case. Since the preceding word for "Ken" is a proper noun ("Kato"), this "Ken" here is supposed to be a Japanese' first name. If, on the contrary, the following word to "Ken" were a proper noun, such "Ken" would probably be an English or American name according to the tendencies in these years.

Ken: I am Kato *Ken*.
健「私は加藤（加東）*健*です．」
Ken「Watashi wa Kato *Ken* desu.」

Meiling: Oh, you are *Ken*. I am Yang Meiling. I am from China.
美玲「ああ，あなたは*健*です．私は陽美玲です．私は中国から来ました．」
Meirin「Aa, anata wa *Ken* desu. Watashi wa Yang Meirin desu. Watashi wa Chuugoku kara ki-mashita.」

Ken: Nice to meet you, Meiling.
                    ....(3)
健「あなたに会えて嬉しいです，美玲．」
Ken「Anata ni aete ureshii desu, Meirin.」

In sentence (3), we have no information to judge whether "Ken" is a Japanese' name

or not within the sentence itself. Here is the case where we apply the information on the sentence before the last. The "Ken" there has been determined to be a Japanese' name because of the reason stated above. Therefore the choice rule for "Ken" in (3)

11 ; ＊健 ; 2 ; E0=Ken ; FT ; DJ1= 「 ; JS2=＊健 ; 2107 ;

checks that the Japanese translation for the last-but-one sentence (JS2) contains "Ken" in kanji character and consequently adopts a Japanese' name "Ken" as the adequate translation. A conversation often goes alternately like ;

A's speech.

B's speech.

A's speech.

B's speech.

So "the sentence before the last" is of much importance in many cases.

4. Which word does 'too' ('also') point out when a noun in the preceding sentence is repeated in the present sentence?

As explained in our previous article [3], what "too" points out in an English sentence is ambiguous, while the Japanese translation will not admit such kind of ambiguity. So we have to decide attaching postposition "mo" whether to the subject or to the object or complement noun in the predicate. Since the same noun as appeared in the preceding sentence can not be attached to by "mo" as explained in the previous article [3], the grammar rule verifies the noun in question to be absent from the preceding sentence before adding "mo" to it.

Contrary to example (3) of the previous article [3], example (4) below has the same noun phrase (E2＝a beaseball fan) in the

predicate of the preceding sentence (BS). Thus the postposition "mo" corresponding to "too" in English is attached to the subject "anata" (you).

Meiling : Are you a baseball fan?
美玲「あなたは野球ファンですか？」
Meirin : Anata wa yakyuu-fan desu ka?

Kenji : Yes, I am. Are you a baseball fan **_too_**? ‥‥‥‥(4)
健二「はい，そうです．あなた**_も_**野球ファンですか？」
Kenji : Hai, sou desu. Anata **_mo_** yakyuu-fan desu ka?

---

3 ; 3216 ; X ; 1 ; T0=v ; T0=? ; FT ; LT ; E3=too ; E2=＊BS ; J0#VR ; J0#DES ; J0#＋か ; J1#＋も ; J2#OJ ; J3#Z ; J1<->J2 ; 7065 ;

appendix

Context checking mechanism in NTT's Japanese-English translation system ALT/JE.

Aroud 1970 when prof. T. Ueno started to develope his English-Japanese machine translation system, a group of researchers in NTT (Nippon Telegram and Telephone Corporation) began to build up their Japanese-English translation system ALT/JE. Both Ueno's system and NTT's system were equipped with the mechanism of checking the context inherited from the preceding sentences.

In an article [5] appeared in 1986, Ueno presented a machine translation system which preserved the Japanese translations of the subject, the verb and the object of the preceding English sentence to use them for the generation of the next translation. On the other hand NTT research group experimented a translation system which used the context data to make clipped pronouns in Japanese sentences recover in their English

translations［6］.

Later, Shibata extended Ueno's system from the view point that each sentence or speech is a sort of man's reaction to the environment, the world which encircles him. In Ueno-Shibata system, various data concerning the preceding and the last-but-one sentences are fully used to the selection of the most adequate translated words in the context and to the generation of the translated sentences.

On the other hand, NTT regretably stopped further developing the context sensitive translation system and since then the US system has become the unique context sensitive machine translation system in the world which is nowadays in incessant progress.

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

［3］ Katsuyuki. SHIBATA: Dependence on Context in case of English-Japanese Machine Translation I-2. Fukuoka University Science Reports, vol.35 No.2, pp.85-88, 2005.
http://www1.fukuoka-u.ac.jp/chosho/Cntxt1-2.html

［4］ Word-Quest Mail, No.269
http://www1.rsp.fukuoka-u.ac.jp/kototoi/2002_6E.html

［5］ Toshio UENO: A Machine Translation System on PC-9801 MS-DOS (2), Context Analysing Routines. PC Magajin, May 1986, pp.76-85.

［6］ Nakaiwa H. and Ikehara S.: Zero Pronoun Resolution in a Japanese to English Machine Translation System using Verbal Semantic Attributes. Transactions of Information Processing Society of Japan, vol.34, No.8, pp.1705-1715 (1993).