

Lexical Transfer based on bilingual signs: Towards interaction during transfer

Jun-ich Tsujii
Kimikazu Fujita

Centre for Computational Linguistics
University of Manchester Institute of Science and Technology
PO Box 88, Manchester M60 1QD, United Kingdom
Email: {tsujii,fujita}@uk.ac.umist.ccl

Abstract

The lexical transfer phase is the most crucial step in MT because most of difficult problems are caused by lexical differences between two languages. In order to treat lexical issues systematically in transfer-based MT systems, we introduce the concept of bilingual-signs which are defined by pairs of equivalent monolingual signs. The bilingual signs not only relate the local linguistic structures of two languages but also play a central role in connecting the linguistic processes of translation with knowledge based inferences. We also show that they can be effectively used to formulate appropriate questions for disambiguating "transfer ambiguities", which is crucial in interactive MT systems.

1. Introduction

'Lexical Transfer' has always been one of the main sources of problems in Machine Translation (MT)[Melby, 1986][Nirenburg, 1988].

Research in transfer-based MT systems has focussed on discovering an appropriate level of linguistic description for translation, at which we can specify 'translation relations' (or transfer rules) in a simple manner. However, lexical differences between languages have caused problems in this attempt. Besides structural changes caused by lexical transfer, selecting appropriate translations of source lexical items has been one of the hardest problems in MT.

Because languages have their own ways of reflecting the structure of the world in their lexicons, and the process of lexicalization is more or less arbitrary, bilingual knowledge about lexical correspondences is highly dependent on language pairs and individual words. We have to prepare a framework in which such idiosyncratic bilingual knowledge about lexical items can be systematically accumulated.

Our approach in this paper follows the general trend in computational linguistics which emphasizes the role of the lexicon in linguistic theory. In particular, our idea of bilingual signs shares a common intuition with [Beaven, 1988] and [Whitelock, 1988]. As with their proposal, we too specify local

structural correspondences between two languages in bilingual lexicons.

Unlike former approaches, however, we explicitly define bilingual signs and use them as predicates in logical formulae (bilingual pivot expressions). Bilingual signs in our framework not only link the local linguistic structures of two languages where the corresponding two monolingual signs appear, but also, by behaving as logical predicates, they connect linguistic-based processes in MT with inference processes. Complicated structural changes, which are often required in translation of remote language pairs like English and Japanese, are captured by logical inferences [Tsujii, 1990].

The framework has the following advantages over conventional methods.

- (i) Reversibility of bilingual dictionaries (lexical transfer rules)
- (ii) Natural interfaces between knowledge-based (inference) processes and MT
- (iii) Ease of paraphrasing using different words (see section 6)

2. Bilingual signs as logical predicates and their definition

The basic idea of bilingual signs is simple. Instead of using predicates corresponding directly to surface words, we use bilingual pairs of lexical items as predicates. That is, we use [RUN:JIKKOOSURU] and [RUN:UN'EISURU] as basic predicates expressing the meanings of *run* in the following sentences.

- (1) The teacher runs the program.
- (2) The teacher runs the company.

Corresponding to the obvious meaning difference of *run* in (1) and (2), we have to use different surface verbs in Japanese, "jikkoosuru" for (1) and "un'eisuru" for (2). The bilingual sign [RUN:JIKKOOSURU] is a predicate which expresses the truth condition which an event should satisfy in order to be described by *run* in English and *jikkoosuru* in Japanese. Note that [RUN:JIKKOOSURU] expresses not only one disambiguated sense of *run* but also one disambi-

guated sense of the Japanese verb *jikkoosuru*¹.

Our system is a conventional transfer based MT system where the monolingual analysis and transfer phases are executed separately. The analysis phase of English produces the following schema of logical formulae (3) as the description of (1). (For simplicity, we ignore articles, quantifiers, etc.)

(3) [[RUN:?1](e) & ARG1(e,x) & ARG2(e,y)
& [TEACHER:?2](x) & [PROGRAM:?3](y)]

(3) is not a logical formula in the ordinary sense but a schema which represents a set of possible formulae. [RUN:?1] is a predicate schema, and by binding the variable '?1' to a specific Japanese verb, we get a specific predicate such as [RUN:JIKKOOSURU], [RUN:UN'EISURU], etc. The transfer phase is taken to be a phase which identifies appropriate predicates in a schema of logical formulae produced by the analysis phase.

As in LFG [Kaplan, 1982], we assume that semantic representations (logical forms) are related lexically with a certain level of linguistic descriptions. Because a bilingual sign is defined by two languages (here English and Japanese), the two relationships of (logical form \leftrightarrow English) and (logical form \leftrightarrow Japanese) are specified in the same place. In order to avoid further complications caused by changes of grammatical functions (passive constructions, etc.), we use thematic role representations as linguistic descriptions in the definitions of bilingual signs.

The following definition shows the predicate [RUN:UN'EISURU] has arity two (arg1 and arg2) and the arguments have sortal restrictions.

(4) (Def-Pred [RUN:UN'EISURU]
{arg1 := [HUMAN:NINGEN] v
[ORGANIZATION:SOSHIKI],
arg2 := [ORGANIZATION:SOSHIKI],
eng := {head := {e-lex := run},
ag1 := <! arg1>,
obj := <! arg2>},
jpn := {head := {j-lex := un'eisuru},
ag1 := <! arg1>,
obj := <! arg2>}})²

This example is rather simple, since local linguistic structures in both languages are the same. That is, the agent and the object in English correspond to the constituents with the same

¹ *jikkoosuru* can be translated into several English verbs including *run*, *carry out*, *execute*, *implement*, *practice*, etc.

² Angle brackets '< >' show a path description and exclamation-mark '!' in the angle brackets means the smallest description block (shown by braces '{ }') which contains the description block in which the '!' appears.

thematic roles. Note that these correspondences are expressed through arg1 and arg2 of the defined predicate. However, many cases have been observed where lexical transfer causes structural changes. It is also the case that objects or events describable by single words in one language are described by phrases or clauses in other languages (see section 3).

We may expect that classes of objects/events which can be expressed by single words in one language correspond to natural classes of objects/events, the classes whose truth conditions are naturally captured by single predicates in logical forms. Therefore, we prepare single bilingual signs for expressing their truth conditions if at least one of the languages has lexical items [Emele, 1990]. That is, we define a single bilingual sign which corresponds to a complex linguistic object in one language, if the other language expresses the same "meaning" by a single word.

As [Sadler, 1990] pointed out, compared with other methods using arbitrary predicates in meaning representation, our method is well-motivated in selecting basic predicates. In fact, the required fineness of distinction of word senses depends highly on the target language (source words are translationally ambiguous [Tsuji, 1988]). We can expect the set of bilingually defined predicates to have appropriate, at least necessary if not sufficient, granularity of the semantic domains for translation of the two given languages.

Furthermore, we can use logical formulae to specify mutual relationships among bilingual signs, which means that we can specify explicitly 'logical' relationships among lexical transfer rules (see section 4).

3. Complex structural changes - complex bilingual signs

The following show how our framework treats structural changes caused by lexical correspondences.

[A] Case changes

The English sentence '*I like him.*' is usually translated into '*Il me plaît.*' in French.

(5) (Def-Pred [LIKE:PLAIRE]
{arg1 := ,
arg2 := ,
eng := {head := {e-lex := like},
ag1 := <! arg1>,
obj := <! arg2>},
fre := {head := {f-lex := plaire},
ag1 := <! arg2>,
obj := <! arg1>}})

In our framework, corresponding case elements in the two languages are linked with each other through the same argument names of bilingual signs.

[B] Lexical inclusions of arguments

A Japanese verb *nuru*, for example, is translated as *paint*, *varnish*, *spread* (bread with butter), *apply* (paint) etc., depending on the material being applied. Some of the English verbs (*paint*, *varnish*, etc.) include the objects (of the Japanese) in their meaning. For example, the structural change between (6a) and (6b) is treated by the definition (7).

- (6a) kabe-ni penki-wo nuru
 [n:wall-location] [n:paint-object] [v]
- (6b) (someone) paints the wall.
 [v] [object]

- (7) (Def-Pred [PAINT:PENKI-WO-NURU])
 {arg1 := ,
 arg2 := ,
 eng := {head := {e-lex := paint}},
 agt := <! arg1>,
 obj := <! arg2>},
 jpn := {head := {j-lex := nuru},
 agt := <! arg1>,
 obj := {head := {j-lex := penki}},
 loc := <! arg2>}})

Note that the Japanese verb *nuru* governs three dependents but one of them is in this definition filled in advance by a specific noun (*penki* – *paint* in English). The definition shows that the phrase *penki-wo nuru* in Japanese corresponds to the English *paint* and that this correspondence defines a predicate as a basic unit of semantic representation.

[C] Head switching

One of the well-known examples is the correspondence between the English verb *like* and the Dutch adverb *graag* (which roughly corresponds to *pleasantly* in English). The same kind of phenomena has often been observed in translation between English and Japanese.

The event expressed by the verb *manage* (in the usage of *manage to do something*) is captured by an adverb *nantoka* ('somehow or other' or 'with great effort' in English) in Japanese. The adverb is used to modify the event expressed as an infinitive clause in English.

The correspondence between (8a) and (8b) is captured by the definition (9).

- (8a) watashi-ga nantoka
 [n:I-subject] [adv:somehow or other]
- ronbun-wo shiage -ta
 [n:paper-object] [v:complete] [tense:past]

(8b) I managed to complete {the/a} paper.

- (9) (Def-Pred [MANAGE:NANTOKA])
 {arg1 := ,
 arg2 := [event:dekigoto],
 eng := {head := {e-lex := manage},
 agt := <! arg1>,
 evt := <! arg2>},
 jpn := {<! arg2>,
 agt := <! arg1>,
 /adv := {head := {j-lex := nantoka}}}}³

In this example, though the adverb *nantoka* is not the head of the Japanese deep case description ('jpn'), it is converted into the predicate [MANAGE:NANTOKA] in the logical formula, and the rest of the 'jpn' description into arg2.

[Kaplan, 1989] proposed two ways of treating such head-switching phenomena, one monolingual and the other bilingual. Our treatment in this paper is basically bilingual in the sense that the non-head construction in Japanese is directly related with the English construction in which the corresponding element is expressed as the head. However, if we deem the logical level of representation a separate, more abstract but mono-lingual level of representation, then our method is quite close to the mono-lingual treatment suggested by [Zajac, 1990]. Our contention is that such an abstract level of representation is hard to justify by purely mono-lingual considerations but only possible by bilingual (or multi-lingual) considerations.

4. Definition of sort hierarchies

Sort-subsort relationships among object-sorts such as '[TEACHER:SENSEI]' is a [HUMAN:NINGEN]', etc. are expressed in conventional logic by implications. However, logical implications express various ontologically different relationships among formulae, which have to be treated differently in translation. Sortal relationships such as these are of special importance in translation, because they give alternative linguistic means of describing the same events/objects (a supersort gives a more vague, less specific description than the subsort). We explicitly indicate that a given implication expresses a sortal relationship, as follows.

³ We introduce a new notation. '{<! arg2>, /adv := (...)}' means that the event/object described by this whole description block minus 'adv:={...}' corresponds to the arg2 of the description block immediately above, and '/adv:={...}' is converted into a predicate at the logical level. Note that our treatment of 'nantoka' is essentially the same as the treatment of 'graag' in the MiMo2 formalism [van Noord, 1990] so that it has the same defect. That is, it cannot cope with cases where more than two words which require 'raising' like 'nantoka' occur at the same level.

(Sort-subsort relationships of event-sorts can also be defined in the same manner).

- (10) (-> SUB:[TEACHER:SENSEI](x)
 SUP:[HUMAN:NINGEN](x)).
 ('->' means logical implication)

(10) shows that, if *x* is describable by *teacher* (or *sensei* in Japanese), the same object can be described by a less accurate word like *human*. We deem the process of selecting an appropriate target expression among possible candidates as the process of locating a expression with the appropriate vagueness level.

The English verb *wear* is a well-known example of a translationally ambiguous word when it is translated into Japanese. It can be translated into several different verbs including *haku* ('wear shoes'), *kaburu* ('wear a hat'), *kakeru* ('wear spectacles'), *kiru* ('wear clothes'), etc., depending on what is worn. While we have a complex expression *mini-tsukeru* (*mi* - body, *ni* particle, *tsukeru* - put on) in Japanese which preserves almost the same vagueness as *wear*, to use this as the translation of *wear* leads to an awkward translation if the material to be worn belongs to a specific sort. *kutsu(shoes)-wo mini-tsukeru*, for example, tends to be understood as "the shoes are worn on a non-standard of the body (not on the feet)".

The predicate [WEAR:MI-NI-TSUKERU] can be defined in a way similar to [PAINT:PENKI-WO-NURU] in (7).

- (11) (Def-Pred [WEAR:MI-NI-TSUKERU]
 {arg1 := [HUMAN:NINGEN],
 arg2 := ,
 eng := {head := {e-lex := wear},
 agt := <! arg1>,
 obj := <! arg2>};
 jpn := {head := {j-lex := tsukeru},
 agt := <! arg1>,
 obj := <! arg2>,
 loc := {head := {j-lex := mi}}}})

The sort-subsort relations between [WEAR:MI-NI-TSUKERU] and [WEAR:HAKU] can be defined as follows.

- (12) (<->>
 SUB:[WEAR:HAKU]
 SUP:[WEAR:MI-NI-TSUKERU]
 CON:ARG2(*self,x*) & [SHOES:KUTSU](x)).

The schema (12) which is specified by '<->>' expresses that

- (i) [WEAR:HAKU] is a subsort of [WEAR:MI-NI-TSUKERU],
 (ii) if an event - *self* - belongs to the sort [WEAR:MI-NI-TSUKERU] and if the argument-2 of the event belongs to the sort

[SHOES:KUTSU], then the event also belongs to [WEAR:HAKU].

All the event-sorts related with *wear* in the above have the same argument structure (arity and role). But this continuity of argument structures through sorts is not necessarily guaranteed. A sort can have multiple supersorts and so the continuity of argument structures from different supersorts may conflict with each other. Furthermore, it is sometimes the case that the arities of events change between a sort and its subsorts. For example, suppose that we have two event sorts [APPLY:NURU] (this event-sort corresponds to the usage of *apply* in *apply glue/paint to ...*) and [PAINT:PENKI-WO-NURU], and that we define the latter as a subsort of the former. Then, one of the arguments in the supersort [APPLY:NURU] is lexically included in the subsort [PAINT:PENKI-WO-NURU] so that these two sorts basically have different arities. The definition of [PAINT:PENKI-WO-NURU] is already given as (7). The definition of [APPLY:NURU] is given as follows.

- (13) (Def-Pred [APPLY:NURU]
 {arg1 := ,
 arg2 := [PAINT:PENKI] v [GLUE:NORI],
 arg3 := ,
 eng := {head := {e-lex := apply-to},
 agt := <! arg1>,
 obj := <! arg2>,
 loc := <! arg3>},
 jpn := {head := {j-lex := nuru}
 agt := <! arg1>,
 obj := <! arg2>,
 loc := <! arg3>}}))

The sort relationship between [APPLY:NURU] and [PAINT:PENKI-WO-NURU] is defined as follows.

- (14) (<->> (<*.ARG2>,<ARG2.ARG3>)
 SUB:[PAINT:PENKI-WO-NURU]
 SUP:[APPLY:NURU]
 CON:ARG2(*self,x*) & [PAINT:PENKI](x))

'<*.ARG2>' and '<ARG2.ARG3>' in this notation mean that the argument-2 in the supersort disappears in the subsort and that the argument-3 in the supersort is mapped to the argument-2 in the subsort. 'ARGi' in the CON-part is taken as referring to the argument structures of the supersort. Unspecified arguments remain unchanged between the sorts.

5. Sketch of the Transfer Phase

The transfer phase is divided into three sub-phases as follows.

- (a) Transforming from thematic role structures of source sentences into schema of logical formulae(like (3))

- (b) Determining logical formulae by descending/ascending sort hierarchies: during this phase, inferences based on knowledge are made, and questions are asked to users, if necessary.
- (c) Transforming from logical formulae to thematic role structures in the target.

All of these steps are performed by referring to the definitions of bilingual signs.

We can index each bilingual sign by the surface word whose 'meaning' is expressed by the sign. Roughly speaking, a word indexing a bilingual sign is either the word which appears as head in the linguistic form definitions or the word which is the value in a feature marked by '/' (like *nantoka* in the example [MANAGE:NANTOKA]).

Step (a) in the above is a rather straightforward process which can be recursively performed through thematic structures. At each recursion level, the system

- (i) identifies the (semantic) head of the level,
- (ii) retrieves the vaguest possible bilingual signs for the head word
- (iii) transforms the local structures governed by the head word according to the definition of the bilingual signs retrieved at (ii).

Because a predicate schema of a word may have several possible vaguest sorts, step (a) produces several formulae which step (b) tries to transform into more appropriate formulae. The processes of descending in sort hierarchies (disambiguation processes necessary for translation) are performed for different predicate schemata simultaneously (for verbs and nouns which are related to each other).

Ascending the hierarchies is also required, because the system has to instantiate all the predicate schemata contained in formula, and constraints imposed by different predicates in a schema of formulae may conflict with each other. It may also happen that there are no corresponding target lexical items for source items. In these cases, the system has to loosen constraints by ascending hierarchies. Therefore, step (b) is a kind of relaxation process which tries to find the most accurate solutions satisfying all constraints. During this process, some general inference mechanisms may be invoked to infer necessary information for navigating in hierarchies and, if necessary, questions will be posed to human users.

[Estival, 1990] also proposed using a partial order of transfer rules to choose preferred translations or prevent less preferred translations from being generated. He assumes that such a partial order of rules can be *automatically* computed in

terms of specificities of conditions on individual transfer rules. We also use a partial order of rules (in our case, lexical transfer rules) to choose translations, but the specificity relationships in our system are concerned with lexical semantics and are not *automatically* computed but defined externally by a human based on his/her bilingual intuition. These externally imposed specificity (sort-subsort) relationships also define possible paraphrasing and are effectively used to disambiguate transfer ambiguities by dialogue.

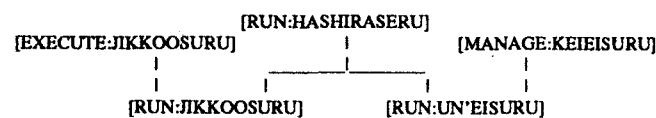
6. Disambiguation of transfer ambiguities by paraphrasing

Because of the explicitness of mutual relationships in the sort hierarchies, we can easily express an event (or object) in diversified ways in both languages. This paraphrasing facility is very useful for forming and posing appropriate questions during the transfer phase to monolingual users of the source language.

Consider the following situation:

(15a) Input sentence: *The teacher runs X.*

(15b) System's knowledge about sorts:



As we have already seen, *run* can be translated into several different verbs in Japanese. Suppose that the sort [RUN:HASHIRASERU] is the least specific sort which *run* can describe. An event of this sort can be directly transformed into Japanese expressions by using *hashiraseru*. However, the direct translation is sometimes awkward if more specific lexical items exist.

The system tries to descend in the hierarchy. In this example, there are two candidates: [RUN:JIKKOOSURU] and [RUN:UN'EISURU]. Three ways of disambiguation by questions are possible: verbalize sort restrictions on arguments directly (ex: (16)), use the other event-sorts which are not shared by both sorts such as (17), and use these two strategies (ex: (18)).

- (16) Is X an organization or a computer program?
- (17) Does the teacher execute X or does the teacher manage X?
- (18) Does the teacher execute X [a program] or does the teacher manage X [an organization]?

7. Conclusion and further discussion

In this paper, we have shown that

- (a) our idea of bilingual signs is useful for representing the relations among lexical transfer rules which in traditional systems

have not been captured explicitly. By using these relationships, we can pose appropriate questions to the user for disambiguation.

- (b) transfer rules which are written in our framework are basically reversible.
- (c) the bilingual signs connect the linguistic forms of two languages and general knowledge about events/objects denoted by them (knowledge about sort hierarchies is the simplest example of this type of knowledge) in a natural way.

In our future research, we have to make it clear to what extent we can treat structural changes by bilingual signs, and on the other hand, to what extent global structural changes beyond the local restructuring by bilingual signs are necessary. We think at present that most of the global structural changes in conventional transfer systems, though necessary for natural translations, actually change the "meanings" of source sentences and should be treated by inference mechanisms external to the "linguistic" processing in translation. Though we only treat the predicates and arguments of bilingual signs, we would have to treat adjuncts as well in order to translate a whole sentence. This is related to how to control the rule application and how to ensure that all the parts of the source structure are processed. The method of formulating questions for disambiguation is still incomplete, though our method seems promising. We have to investigate what sorts of paraphrasing are really helpful for making bilingual ambiguities obvious to monolingual users.

Acknowledgements

This work is supported partly by the research contract with ATR (Advanced Telecommunication Research Lab.) in Japan. We are grateful to the members of the research group at CCL, UMIST (Dr.J.Carrol, Mr.J.Lindop, Dr.M.Hirai, Mr.J.Phillips, Dr.H.Somers and Dr.K.Yoshimura) for their valuable discussions.

References

- [Alshawi, 1989]: Alshawi, H. and van Eijck, J.: Logical Forms in the Core Language Engine, in Proc. of 27th ACL, Vancouver, 1989.
- [Beaven, 1988]: Beaven, J. and Whitelock, P.: Machine Translation Using Isomorphic UCGs, in Proc. of Coling-88, Budapest, 1988.
- [Emele, 1990]: Emele, M., Heid, U., Momma, S. and Zajac, R.: Organizing linguistic knowledge for multilingual generation: in Proc. of Coling-90, Helsinki, 1990.
- [Estival, 1990]: Estival, D., Ballim, A., Russell, G. and Warwick, S.: A Syntax and Semantics for Feature-Structure Transfer, in Proc. of The 3rd International Conference on Theoretical and Methodological Issues in Machine Translation, Austin, 1990.
- [Kaplan, 1982]: Kaplan, R. and Bresnan, J.: Lexical Functional Grammar: a formal system for grammatical representation, in Joan Bresnan(ed.), The mental representation of grammatical relations, MIT Press, 1982
- [Kaplan, 1989]: Kaplan, R., Netter, K., Wedekind, J. and Zaenan, A.: Translations by structural correspondences, in Proc. of 4th European ACL Conference, Manchester, 1989.
- [Melby, 1986]: Melby, A.K.: Lexical Transfer: Missing Element in Linguistic Theories, in Proc. of Coling 86, Bonn, 1986.
- [Nirenburg, 1988]: Nirenburg, S. and Nirenburg, I.: Framework for Lexical Selection in Natural Language Generation, in Proc. of Coling 88, Budapest, 1988.
- [Sadler, 1990]: Sadler, V.: Working with Analogical Semantics, Distributed Language Translation, Foris, 1990.
- [Tsujii, 1986]: Tsujii, J.: Future Directions of MT, in Proc. of Coling 86, in Bonn, 1986.
- [Tsujii, 1988]: Tsujii, J. and Nagao, M. : Dialogue Translation vs. Text Translation, in Proc. of Coling 88, Budapest, 1988.
- [Tsujii, 1990]: Tsujii, J., Fujita, K. : Lexical Transfer based on bilingual signs, in Issues in Dialogue Machine Translation (CCL report no. 90/5), 1990.
- [van Noord, 1990]: van Noord, G., Dorrepaal, J. et.al.: The MiMo2 Research System, in Proc. of The 3rd International Conference on Theoretical and Methodological Issues in Machine Translation, Austin, 1990.
- [Whitelock, 1988]: Whitelock, P.: The organization of a bilingual lexicon, DAI Working Paper, Dept. of Artificial Intelligence, Univ. of Edinburgh, 1988.
- [Zajac, 1990]: Zajac, R.: A relational approach to translation, in Proc. of The 3rd International Conference on Theoretical and Methodological Issues in Machine Translation, Austin, 1990.