LFG and Translation

Louisa Sadler Ian Crookston Doug Arnold
Andy Way
University of Essex
Colchester C04 3SQ UK *

Introduction

Recent work in LFG uses the notion of projection to refer to linguistically relevant mappings between levels, whether these mappings are direct or involve function composition ([2],[3],[4],[6],[7]). Work in this spirit by Kaplan et al [7] defines two translation functions τ (between f-structures) and τ' (between semantic structures). By means of these functions, one can 'co-describe' elements of source and target f-structures and s-structures respectively. This approach to translation departs from the classical transfer model in a number of respects and seems to offer a number of advantages over that model. Achieving translation can be thought of in terms of specifying and resolving a set of constraints on target structures, constraints which are expressed by means of the τ and τ' functions. In common with much contemporary work in computational linguistics, the approach is description-based rather than constructive, and it preserves systematicity without imposing the constraint of compositionality [2]. LFG co-description offers the particular advantage of modularity as a formalism for MT. That is, unlike transfer systems, it does not conflate all translationally relevant information into a single, linguistically hybrid level of representation and yet still allows information from different linguistic levels of representation to interact to constrain the translation relation, by function composition.

The formalism can handle a number of cases of 'hard' translation nicely, but fares less well with the class of cases discussed here. An important feature of this approach is that translation equations are added to source lexical entries and c-structure annotations alongside monolingual equations. This paper shows how picking out the correct units for translation in a set of cases is at least as difficult as in transfer systems. A final section discusses some natural extensions to the formalism intended to capture the specificity of the translation function in comparison to the monolingual mappings. One

* This research has been carried out within the British Group of the EUROTRA project, jointly funded by the Commission of the European Communities and the United Kingdom's Department of Trade and Industry.

of the proposals extends the description language to include the priority union operator. We assume basic familiarity with LFG (see [5] and [11] for a basic introduction) and refer the reader to [7] for further details of the LFG translation functions.

Section One

We assume that the purely linguistic component of an MT system should produce the set of all and only the possible translations (Landsbergen [8]). A source word which is lexically ambiguous (from the point of view of the target language) will produce two translations. For example, English *river* will translate as *fleuve* and *rivière* and Dutch *bank* as *bench* or *settee.* In these cases, we might enclose the alternative *τ* equations in disjunctive braces in the source lexicon. In such cases, the solution algorithm must produce two candidate structures, differing only in the translation of the ambiguous word.

For *bank* the two equations would be

$$\{(\tau \uparrow \text{PRED FN}) = \text{bench}, (\tau \uparrow \text{PRED FN}) = \text{settee}\}$$

A device with contextual and world knowledge has to be assumed to choose between such competing translations.

Satisfying the requirement that only possible translations are produced is problematic where the translation of a lexical head is conditioned in some way by its dependents.

We concentrate on this problem as illustrated in (1):

(1)
know how ⇒ savoir
commit suicide ⇒ se suicider
not until ⇒ erst
until now ⇒ jusqu'ici
too much ⇒ trop

In these cases the source language uses a head + dependent combination while the target translation fuses these two elements. We use the term incorporation to refer to this phenomenon.

Consider the following:

commit a crime ⇒ commettre un crime
commit suicide ⇒ se suicider/*commettre le suicide

The regular or general translation equations for *commit* are:

(2)

$$(\tau \uparrow \text{PRED FN}) = \text{commettre}$$
$$\tau(\uparrow \text{SUBJ}) = (\tau \uparrow \text{SUBJ})$$
$$\tau(\uparrow \text{OBJ}) = (\tau \uparrow \text{OBJ})$$

stating that the target f-structure has the PRED *commettre,* and that the target f-structure SUBJ and OBJ slots are filled by the translations of the source SUBJ and OBJ respectively.

In order to translate *commit suicide* as *se suicider,* we might add the following set of equations:

(3)

$$\text{a } (\uparrow \text{OBJ PRED}) =_c \text{suicide}$$
$$\text{b } (\tau \uparrow \text{PRED FN}) = \text{se suicider}$$
$$\text{c } \tau(\uparrow \text{SUBJ}) = (\tau \uparrow \text{SUBJ})$$

The constraining equation (3)a is intended to limit the applicability of this set in the obvious manner, permitting us to disjoin this set with the set in (2) above.

Take the problematic source f-structure:

$$(4)\quad \begin{bmatrix} \text{PRED} & \text{commit} \\ \text{SUBJ} & f_1[\ldots] \\ \text{OBJ} & f_2\,[\text{ PRED} \quad \text{suicide }] \end{bmatrix}$$

with the entry for *suicide* as in (5). The regular set gives us (6):

$$(5)\quad (\tau \uparrow \text{PRED FN}) = \text{suicide}$$

$$(6)\quad \begin{bmatrix} \text{PRED} & \text{commettre} \\ \text{SUBJ} & \tau f_1\,[\ldots] \\ \text{OBJ} & \tau f_2\,[\text{ PRED} \quad \text{suicide }] \end{bmatrix}$$

Despite expressing the dependencies between *commit* and *suicide,* we have not suppressed the regular translation.

Now consider the irregular translation, using the exceptional equations for *commit* and the translation of *suicide:*

$$(7a)\quad \begin{bmatrix} \text{PRED} & \text{se suicider} \\ \text{SUBJ} & \tau f_1[\ldots] \end{bmatrix}$$

(7b) τf_2 [**PRED** suicide]

Target wellformedness conditions (completeness and coherence) would correctly prevent the unification of the structures in (7). But notice then that f_2 has effectively remained untranslated. In order to describe a complete target f-structure in which f_2 is translated we now need to make the translation of *suicide* optional:

(8)

suicide: $\{(\tau \uparrow \text{PRED FN}) = \text{suicide}$ [regular]

$(\tau\uparrow) = \text{nil}$ [irregular] $\}$

This is unreasonable and linguistically implausible. It is implausible to state that the translation of a particular lexical item is nil without stipulating the environment under which this state of affairs is true, i.e. there is nothing in the second equation above to say that it is part of the translation information for *commit suicide* ,and nothing else. In any case, what we want to say (see section Three) is that se *suicider* is a translation of *commit suicide* as a whole. This is not achieved by the entry in (8).

Furthermore adding 'nil' equations as in (8) is dangerous, in that it will produce null translations elsewhere, perhaps producing grammatical output which is not translationally equivalent to the source. [1]

Of course, even with the irregular set amended thus, we pass over as possible translations the pair:

commettre le suicide
se suicider

This may be considered reasonable, depending on our view of what constitutes 'possible translation'. Note that phrases like *rolling staircase* (for *escalier roulant)* are perfectly acceptable in the target language, just not the right translation. This problem seems to be tractable only with difficulty in current 'linguistic' MT (see [10] for one approach) and should perhaps be considered as a separate problem.

[1] An attempt to specify in the lexical entry for the governed element exactly when no translation is required involves using

toy:
$(\tau \uparrow) = \emptyset$
$(\phi(M(M(*)))\ \text{PRED} =_c \text{library}$

which reintroduces the use of the closure of the mother relation in the domain of ϕ, which has been rejected in the treatment of unbounded dependencies in LFG. Such a move might have serious consequences for the inspectability and accuracy of translation specifications.

Section Two

While target well-formedness conditions could play some role in restricting the output of τ to grammatically possible translations, as discussed above, they cannot constitute a general solution, since incorporation may involve non-subcategorisable elements (ADJUNCTS). Likewise, we cannot make translations optional in general, as in (8), for then we run the risk of leaving the source adjuncts untranslated. For example:

(9)
toy library ⇒ ludothèque
pierre tombale ⇒ gravestone
épine dorsale ⇒ backbone
bring together ⇒ rapprocher
aller en flottant ⇒ float
plante grasse ⇒ succulent
plante grimpante ⇒ creeper

Following the model of (3) we might assume (11) as the regular τ equation for *toy* and hasard (10) as the 'special' τ case for *library:*

(10)
library:
$[\text{PRED} = \text{toy}] \in_c \uparrow\text{ADJ}$
$(\tau \uparrow \text{PRED FN}) = \text{ludothèque}$

(11)
toy:
$(\tau \uparrow \text{PRED FN}) = \text{miniature}$

There are a number of problems with this.

Note that τ equations for translating adjuncts are annotated to the c-structure rule, as in (12).

(12)

$$\begin{array}{lll} \text{NP} \rightarrow & \text{AP*} & \text{N} \\ & \downarrow \in \uparrow \text{ADJ} & \\ & (\tau \downarrow) \in (\tau \uparrow \text{ADJ}) & \end{array}$$

Making this annotation optional, or disjoining the τ equation on *toy* in (11) with an equation translating it as 'nil', as in section One, would lead to a general problem of non-translation of adjuncts. The alternative will produce target f-structures corresponding to *ludothèque miniature* and so on. The problem, then, is specifying just which adjuncts are to remain untranslated in which contexts.

Furthermore there are serious problems with (10). LFG is formulated in such a way as to exclude reference to adjuncts (members of the set-valued ADJ attribute) in lexical entries:

> "... since there is no notation for subsequently referring to particular members of that set (i.e. the set of adjuncts), there is no way that adjuncts can be restricted by lexical schemata associated with the predicate ... Since reference to the adjunct via the 'down arrow' is not possible from other places in the string, our formal system makes adjuncts naturally context-free." ([5]:216)

But we need to refer to the adjunct *toy* in (10), in order to give context for the translation *ludothèque.* In order to do so, we have been forced to introduce the membership constraint shown in (10). Some revision to allow one to refer to a particular member of the set of adjuncts is therefore necessary, and will be assumed in the following section.

Section Three

The problem is that lexical dependencies such as those discussed in sections One and Two cannot be simply captured by τ or τ' in the framework of [7], since they are not and should not be explicitly recognised as special on monolingual grounds.

It is useful here to compare this model with constructive or representation based transfer systems. In such systems, transfer rules are derivational, providing instructions for the construction of target representations, roles which match and may take apart source representations. In LFG MT representations play little or no part—the piecewise mapping function T instead involves equations which attach to source lexical items and to source c-structure rules, not to representations. So, where we want to say that the representation or representation subpart which has *commit* as head and *suicide* as object translates as *se suicider,* we cannot. We are forced to try to get the same effect by saying things about *commit* and *suicide* in isolation. Of course, the splintering of translation information across lexical items gravely affects the transparency of grammars.

Contrast this with a typical rule from a representation-based system (in this case MiMo [1]), which extracts from the source representation and inserts into the target representation those parts which we wish to translate as a unit:

(13)
!commit.[arg2=!suicide] ⇔ !se-suicider

(13) specifies that *commit* and *suicide* together translate as *se suicider* and

vice versa. In addition, it is possible to organise a transfer component so that such rules suppress (or override) the 'regular' translation pair *commit* ⇔ *commettre.*

In the rest of this section, we sketch out a revision of the notion and role of transfer dictionary in LFG in order to capture these dependencies without tuning our f-structures which must remain justifiable on purely monolingual grounds. It is worth noting that under our proposal, the structure of the transfer lexicon, in terms of its units of organisation, differs from that of the monolingual lexicon, setting the τ projection apart from the monolingual projections. We believe that this is correct.

The fundamental problem lies in the nature of the 'transfer lexicon' assumed in the model in [7]. Transfer or translation equations come from either the source c-structure annotations or from a transfer lexicon. Problematically for the present data, this lexicon is organised according to the source language grammar—that is, it contains τ equations for the units of the source lexicon. It is thus most naturally thought of as simply a further set of equations in the monolingual source lexicon - just another projection. But the data we are concerned with show that the units for translation are not exactly co-extensive with the units for monolingual analysis. The transfer lexicon should reflect this fact. That is, there must be entries in the transfer lexicon for each of these cases. In what follows, we assume for clarity that the 'regular' transfer equations (2) and (5) reside in the monolingual dictionary along with other projection information. We introduce in addition a true transfer lexicon, which crucially views the source f-structure through the perspective of the target language. Thus it is organised around translationally relevant units.

We illustrate this transfer lexicon with *commit* (of *suicide*). As a first approximation:

(14)
commit:
a (↑ OBJ PRED) $=_c$ suicide
b (τ ↑ PRED FN) = se suicider
c τ (↑ SUBJ) = (τ ↑ SUBJ)
d τ (↑ OBJ) = Øor
d τ (↑ OBJ) = τ ↑

The regular τ equation from the lexical entry for *suicide* of course assigns the FN *suicide* in the translation. This information is inconsistent with d) and must be overridden. Rather than change or add to the regular equation for *suicide,* we assume that priority union is necessary to prefer the information in the transfer lexicon over any inconsistent information from the regular τ equations. Thus an entry like (14) will bear the priority operator (see [6] for some brief discussion). Of course, such operators must be treated with care.

As before, we produce *commettre le suicide* and *se suicider* as candidate translations. But we maintain the transparency of our lexical entries—most importantly, we do not alter the regular entry for *suicide*. It is easy to see that some mechanism to prefer target f-structures produced from the 'special' t-lexicon can be added, if required.

Consideration of cases where non-subcategorised elements must be incorporated or fused in translation shows that the picture is slightly more complicated. In section Two these cases were problematic because we could not constrain the null translation or the regular translation of the adjunct appropriately, producing for *toy library* both *ludothèque* and *ludothèque miniature,* and for *toy car* both *voiture miniature* and *voiture* respectively. Our special entry for such combinations of head and adjunct or modifier must ensure this.

The problem is that of specifying which adjunct serves as context and is to be 'untranslated', for we need to specify not a path in the source, but a value (roughly $(\tau \uparrow \text{ADJ PRED toy}) = \text{NIL}$ or $\tau (\uparrow \text{ADJ PRED toy PRED FN}) = \text{NIL}$). Such an equation would be inconsistent with the regular translation of *toy* as *miniature* in the normal way, and could again be treated by priority union. It is clear that the current impossibility of referring to members of the ADJUNCT set is problematic and some extension of the notation may be required.

However this could be simulated by structuring the transfer lexicon further to allow conjoined entries:

(15)
[[*library:* $(\tau \uparrow \text{PRED FN}) = \text{ludothèque}$]
$\wedge$ [*toy:* $(\tau \uparrow) = \text{NIL}$]]

The use of conjunctions of entries as in (15) can of course be extended to the subcategorised cases (14), obviating the need for priority union with the regular equations in that case too.

Conclusion

We have discussed some cases which seem to be problematic for the τ projection as presented in [7]. We have suggested that the problem arises because the transfer lexicon is organised solely around the lexical units of the source language, thus making it impossible to state context for some translations. In section Three we have presented some extensions to the framework to partially overcome these problems in what is otherwise a very nice translation formalism.

Bibliography

[1]
Arnold, D.J., S. Krauwer, L. Sadler & L. des Tombe (1988) "Relaxed Compositionality in Machine Translation", **Proceedings of the Second Conference on Theoretical and Methodological Issues in Machine Translation**, CMU, Pittsburgh.

[2]
Halvorsen, P-K (1988) "Algorithms for Semantic Interpretation", presented at Workshop on Computational Linguistics and Formal Semantics, Lugano, Switzerland.

[3]
Halvorsen, P-K (1988) "Algorithms for Semantic Interpretation", presented at Workshop on Computational Linguistics and Formal Semantics, Lugano, Switzerland.

[4]
Halvorsen, P-K & Kaplan, R.M. (1988) "Projections and Semantic Description in Lexical-Functional Grammar" presented at the International Conference on Fifth Generation Computer Systems, Tokyo, Japan.

[5]
Kaplan, R. & J. Bresnan (1982) "Lexical-Functional Grammar : A Formal System for Grammatical Representation" in J.Bresnan (ed) **The Mental Representation of Grammatical Relations**, MIT Press, Cambridge, Mass.

[6]
Kaplan, R. (1987) "Three seductions of computational psycholinguistics", in P.Whitelock, M.M.Wood, H.L.Somers, R.Johnson & P.Bennett (eds) **Linguistic Theory and Computer Applications**, Academic Press, London, pp 149-88.

[7]
Kaplan, R., K. Netter, J. Wedekind & A. Zaenan (1989) "Translation by Structural Correspondences", **Proceedings of the 4th EACL**, UMIST, Manchester, pp 272-81.

[8]
Landsbergen, J. (1987) "Montague Grammar and Machine Translation" in P.Whitelock et al (eds), pp 113-148.

[9]
Sadler, L. et al (1989) "Co-description, projection and 'difficult' translation", Working Papers in Language Processing No. 8, University of Essex.

[10]
Schubert, K. (1987) **Metataxis: Contrastive Dependency Syntax for**

Machine Translation, FORIS, Dordrecht, Holland.

[11]
Sells, P. (1985) **Lectures on Contemporary Syntactic Theories,** CSLI, Stanford University, California.