

Codescription and Transfer

Louisa Sadler, University of Essex, UK.
SALT Workshop, July 1990

Introduction

Recent work in LFG [1] uses the notion of projection to refer to linguistically relevant mappings or correspondences between levels of description, whether these mappings are direct or involve function composition. ([2],[3],[4],[5]). Mapping functions such as phi (from C to F structure) and sigma (from C to Semantic structure) are familiar from the LFG literature. In [5], Kaplan et al define two translation functions tau (between f-structures) and tau' (between semantic structures). By means of these functions, one can 'co-describe' elements of source and target f-structures and s-structures respectively. 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 tau and tau' functions.

The formalism is extremely flexible in permitting a wide variety of source-target correspondences to be expressed: tau and phi can be composed, as can tau' and sigma. Simply put, the approach allows for equations specifying translations to be added to lexical entries and (source language) c-structure rules. For example, (1)

$$(1) \tau (\wedge \text{SUBJ}) = (\tau \wedge \text{SUBJ})$$

composes tau and phi, equating the tau of the SUBJ f-structure (here on the lefthand side of the equality) with the SUBJ attribute of the tau of the mother's f-structure (here on the righthand side of the equality). Thus (1) can be thought of as saying that the translation of the value of the SUBJ slot in a source f-structure fills the SUBJ slot in the f-structure which is the translation of that source f-structure.

This approach to translation departs from the classical transfer model in a number of respects. Firstly, it is description-based rather than constructive or derivational in approach. Secondly, it preserves systematicity without imposing the constraint of compositionality ([2],[4]). Thirdly, LFG co-description as a formalism for MT appears to offer the particular advantage of modularity. That is, unlike transfer systems it does not have to 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. Fourthly, it should be noted that the model of translation presented in [5] is not bidirectional.

I hope to indicate that there are at least two sorts of linguistic case which arise in MT which are problematic for this approach. This is a report on joint work, and a fuller presentation of these problems is given in [6] and [7].

The first problem involves the translational phenomenon of head-switching (illustrated in (2) below), which is much discussed in the MT literature. The problem is that the treatments proposed in [5] do not extend straightforwardly to all cases of this type because of a problem in integrating 'special' and 'regular' equations. The second problem concerns the difficulty of picking out the correct units for translation in this approach.

Even closely related languages may differ in their choice of syntactic head in translationally equivalent constructions.

- (2)
- (a) Jan zwemt graag/toevallig
John likes to/happens to swim
 - (b) John has just arrived
Jean vient d'arriver
 - (c) Ik denk dat Jan graag zwemt
I think that John likes to swim
 - (d) Ich glaube dass Peter gern schwimmt
I believe that Peter likes to swim

In these examples, an adverb in (a) (which is not the syntactic head of the construction) corresponds translationally to a verb in (b). Such cases typically involve the writing of complex rules in a transfer system.

Kaplan et al [5] sketch two alternative approaches to head-switching. The first involves performing head-switching "monolingually" by adopting an f-structure which is right for translation into a particular language (e.g. by treating the Dutch adverbs in 2 (a) and (c) as syntactic heads of their constructions at f-structure).

This approach is illustrated in (4). Note that the PS rules are based on those in [5] and are purely illustrative.

- (3a) Jan zwemt graag
(3b) John likes to swim
- (4) s -> NP VP ADVP*
 (SUBJ)=v ^=v ^=(v ARG)

graag: ADV (^PRED) = graag<ARG>
(tau ^ PRED FN) = like
tau (^ ARG) = (tau ^ XCOMP)

zwemmen:V (^PRED) = zwemmen<SUBJ>
(tau ^ PRED FN) = swim
tau(^ SUBJ) = (tau ^ SUBJ)

The essential characteristic of this is the f-description on ADVP

which indicates that the f-structure associated with the mother (S) node is the value of the ARG attribute of the f-structure associated with the ADVP node. The tau equations relate (5) to (6) (information such as the necessary functional control equation for like comes from the English lexicon).

```
(5)  [PRED      graag<ARG>
      ARG       fl[ PRED      zwemmen<SUBJ>
                   SUBJ      [PRED      jan]]]
```

```
(6)  [PRED  like<SUBJ,XCOMP>
      SUBJ   [ ]
      XCOMP  [PRED  swim<SUBJ>
              SUBJ  [PRED John]]]
```

The problem arises in cases like (2c) in which the adverb occurs within a complement clause. The normal f-description on embedded S nodes ($\hat{\text{COMP}} = v$), specifies that the f-structure associated with the embedded S node (fl in (5) above) is the COMP of the dominating PRED. The result of combining this equation with those in (4) above is a doubly rooted DAG, which is not the required structure. The problem, then, is that the "special" equation in (4) does not combine correctly with the "regular" equation for the embedded S node.

The authors note an alternative treatment in which the adverbs are f-structure SADJs, with head-switching occurring between source and target f-structures.

In (7) the tau annotation to ADVP states that the tau of the mother f-structure is the XCOMP of the tau of the SADJ slot.

```
(7)  S -> NP          VP  ADVP
      (^SUBJ)= v   ^=v   (^SADJ)=v
                        (tau ^SADJ) XCOMP = tau ^
graag: ADV (^PRED) = graag
(tau ^ PRED FN) = like

zwemmen:V      (^PRED)   = zwemmen<SUBJ>
(tau ^ PRED FN) = swim
tau (^ SUBJ) = (tau ^ SUBJ)
```

```
(8a)  f3[PRED zwemmen<SUBJ>
      SUBJ f4 [PRED Peter]
      SADJ f5 [PRED graag]]
```

```
(8b) tf5[PRED    like<SUBJ,XCOMP>
      SUBJ      [tf4]
      XCOMP    tf3[PRED    swim<SUBJ>
                  SUBJ    tf4 [PRED Peter]]]
```

Note that the Dutch f-structure is rooted in f3 and the English in tf5.

In this approach, examples like (2c-d) are again problematic. The lexical entry for denken would contain 'regular' translation equations:

```
(9) denken: V (^PRED) = denken<SUBJ,COMP>
      (tau ^ PRED FN) = think
      tau (^ SUBJ) = (tau ^ SUBJ)
      tau (^ COMP) = (tau ^ COMP)
```

This specifies that tau of the mother's f-structure's COMP slot is the COMP of the tau of the mother's f-structure. However, the annotation on ADVP in (7) requires the tau of the same piece of f-structure to be the XCOMP of the tau of the SADJ slot. What results is a doubly rooted, reentrant in tau f3:

```
tau (^ COMP) = (tau ^ COMP)      tau f3 = tau f1 COMP
(tau ^) = tau (^ SADJ XCOMP)    tau f3 = tau f5 XCOMP
```

The problem is to achieve simple general statements of the correspondences which cover exceptional cases which interact correctly with the equations for the "regular" correspondences.

We now turn very briefly to some data which show that the units for translation are not co-extensive with the units for monolingual analysis (see [7] for a fuller discussion). This second problem concerns the phenomenon of dependent incorporation, illustrated below:

```
(10) know how -> savoir
      commit suicide -> se suicider
      not until -> erst
```

Our starting point is the following observation:

```
commit a crime -> commettre un crime
commit suicide -> se suicider/*commettre le suicide
```

The 'regular' translation equations for commit are:

```
(11) (tau ^ PRED FN) = commettre
      tau ( ^ SUBJ) = (tau ^ SUBJ )
      tau ( ^ OBJ) = (tau ^ 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 for 'exceptional' commit:

(12)
a (\wedge OBJ PRED) =c suicide
b ($\tau \wedge$ PRED FN) = se suicider
c $\tau \wedge$ SUBJ = ($\tau \wedge$ SUBJ)

(12a) is intended to limit the applicability of this set to cases where the OBJ PRED is suicide and (12) is disjoined with the set in (11). The translation equation for suicide is:

(13) ($\tau \wedge$ PRED FN) = suicide

The source f-structure (14) and the equations in (11) produce (15):

(14)
[PRED commit
 SUBJ f1[...]
 OBJ f2[PRED suicide]]

(15)
[PRED commettre
 SUBJ τ f1[...]
 OBJ τ f2[PRED suicide]]

Using the exceptional equations for `commit` (12) and the translation of suicide produces:

(16a) [PRED se suicider
 SUBJ τ f1[...]]

(16b) τ f2 [PRED suicide]

Of course, these structures cannot be unified. In order to describe a complete target f-structure in which f2 is translated we need also to make the translation of suicide optional:

(17)
suicide: { ($\tau \wedge$ PRED FN) = suicide [regular]
 ($\tau \wedge$) = nil } [irregular]

The result is that we still incorrectly produce both `commettre le suicide` and `se suicider` as possible translations. Moreover (17) is unreasonable and linguistically implausible, since no environment is stated for the 'irregular' translation. It is dangerous since it will produce null translations elsewhere. Notice

that even (18) cannot prevent the production of (16b) and of commettre le suicide.

(18) commit
a (^ OBJ PRED) =c suicide
b (tau ^ PRED FN) = se suicider
c tau (^ SUBJ) = (tau ^ SUBJ)
d tau (^ OBJ) = nil or
d tau (^ OBJ) = tau ^

Finally, while it is perhaps in principle possible to use constraining equations in the monolingual dictionary to help for these subcategorised cases, as above, this is not a general solution, since incorporation may involve non-subcategorisable elements (ADJUNCTS):

(19) pierre tombale -> gravestone
épine dorsale -> backbone
bring together -> rapprocher
aller en flottant -> float
plante grasse -> succulent
plante grimpante -> creeper

ADJUNCT is a set-valued feature in f-structure, and it is impossible to pick out the relevant ADJUNCT and to ensure that it translates as 'nil' without producing a general problem of non-translation of adjuncts. The problem, then, is that of specifying just which adjuncts are to remain untranslated in which contexts.

BIBLIOGRAPHY

[1] 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.

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

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

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

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

[6] Sadler, L., I. Crookston and A. Way (1989) "Co-description, projection and 'difficult' translation", Working Papers in Language Processing No. 8, Department of Language and Linguistics, University of Essex.

[7] Sadler, L., I. Crookston, D. Arnold and A. Way (1990) "LFG and Translation", in Third International Conference on Theoretical and Methodological Issues in Machine Translation 11-13 June 1990, Linguistics Research Center, Austin, Texas.