Abstract:

This paper tries to investigate valency theory as a linguistic tool in machine translation. There are three main areas in which major questions arise:

1) Valency theory itself. I sketch a valency theory in linguistic terms which includes the discussion of the nature of dependency representation as an interface for semantic description.

2) The dependency representation in the translation process. I try to sketch the different roles of dependency representation in analysis and generation.

3) The implementation of valency theory in an MT-system. I give a few examples for how a valency description could be implemented in the EUROTRA formalism.

0. Introduction

This paper tries to apply a piece of established linguistics, known as valency theory, to the problem of machine translation. As such, it is meant to fit into the forthcoming EUROTRA MT system, though it does not deal with EUROTRA problems specifically.

There are two aspects which play a role in the building of an MT-system:

(1) the development of linguistic specifications

(2) the development of a formalism which allows for the implementation of linguistic results.

This paper mainly deals with the first aspect and thus relates not only to EUROTRA but also at least to all stratificational systems, i.e. systems that break up the translation process into a sequence of simpler translation processes. Furthermore it relates at least to any system which uses dependency/valency information as e.g. EUFG does with its functional structure.

In EUROTRA, the level where information about dependency/valency is used is the ERS (Eurotra-Relational-Structure) which lies between the constituent structure (ECS) and the semantic interface structure (IS).

So, in EUROTRA terms I try here to give a kind of ERS-definition in the language of empirical linguistics without touching the formalism itself.

The investigation divides into three parts:

(1) the sketch of a valency theory which comprises the following points:

- the informal definition of the concepts valency, complement, and adjunct, thereby trying to give a definition which holds for verbs, adjectives, nouns, and prepositions,

- the operationalization of the complement adjunct distinction derived from this definition,

- the classification of the complements, a subclassification of the complement classes (C-classes) according to their syntactic realization, and the determination of the relevant sentence patterns,

- a short discussion of the relation of the dependency level to constituency level and from dependency level to semantic representation.

(2) the application of the linguistic specifications to the problem of MT which has to investigate the role of dependency representation (D-representation) (or rather the role of the transitions to D-representation, since levels do not "play a role") in analysis and generation.

(3) the way the gained linguistic information can be implemented according to the EUROTRA formalism.

1. Outline of a Syntactic Valency Theory

1.1 Definition of Concepts

The syntactic valency of an element of a word class (a nonterminal category) is its property to bind a certain number and a certain kind of syntagms.

Those valency-bound syntagms are the complements. The syntagms which are not valency-bound are the adjuncts. All syntagms which are obligatory in a syntactical sense are valency-bound, i.e. the sentence must contain that item otherwise it would not be complete. Besides the obligatory syntagms the wordsubclass specific elements are valency-bound (Engel 1982). Wordsubclass specific elements are those syntagms which can occur only with elements of wordsubclasses.

1.2 Operationalization of valency

This point usually forms the major part of each paper on valency theory and mostly consists of a discussion of various proposed tests in the literature on valency theory showing their deficiencies and coming up with a new proposal.
The discussion of proposed tests will be left out in this paper and the proposal for an operationalization that will be made consists of an addition of a test which has already been established and which could be called "the free addibility test". This test will be refined here and is thus a "modified addibility test." It will be explained mainly by applying it to verb valency but is also meant to apply to word-classes other than verbs.

In order to operationalize the term "obligatory complement" the elimination test is applied. The syntagmas in a given sentence is obligatory if it cannot be eliminated without rendering the sentence ungrammatical. In German, only verbs, adjectives and prepositions can have obligatory complements.

In order to operationalize subclass specificity, a modified addibility test is applied. Throughout the history of the theory of valency, many tests have been proposed. The one proposed here is considered the most adequate as there is an obvious relation between definition of concept and operationalization. It also has the property of being partially applicable interlinguistically, which is not unimportant for German.

The "free addibility test" consists of checking which syntagmas occur together. If there are no restrictions for a syntagma, i.e. if it can be freely added, it is an adjunct and not valency bound. This test, however, cannot handle the following problems. The respective explorations lead to a refinement of the test and thus to the "modified addibility test".

a) There are restrictions in combining temporal adjuncts with tenses of the verb.

(1) Er kam morgen.
(He came tomorrow.)

b) There are restrictions in combining certain elements of adjectival classes with certain subjects.

(2) Dec Year nicht absichtlich um Bang.
(He tries not deliberately on the slope.)

In (a) and (b), it can be proved that the restrictions do not come from the verb but are caused by other reasons. It is possible to force correct sentences containing both the verb and the adjunct. Therefore, there is no principled impossibility of their combination.

(3) Er kom: morgen.
(He will come tomorrow.)
Fritz steht absichtlich am Abhang.
(Fritz stands deliberately on the slope.)

c) More problematic is the case where elements of subclasses of adjective classes are in fact not combinable with a verb. In

(4) Fritz atmet gemütlich.
(He breathes cosily.)

the restriction comes in fact from the verb. The syntagma "gemütlich" which is denoted as an adjunct would fulfill the requirement of being subclass specific. Therefore, the addibility test has to be restricted.

(5) Fritz atmet schwer.
(He breathes heavily.)

As (5) is correct and "schwer" belongs to the same class of adjectives, i.e. adjectives of manner, the addibility test will have to be restricted to classes of syntagmas. If elements from adjectival classes are freely combinable, the addibility test is considered as fulfilled.

d) However, there are certain problematic cases.

(6a) *Dez am mit einem Hämmer.
(*He is talking with a hammer.)

(6b) Ich feste die Holz morgen.
(Who can it not until tomorrow.)

Instrumental adjectives and certain temporal adjectives can scarcely be described as being freely combinable as they only combine with action verbs and so fulfill the criterion of subclass specificity. On the other hand, they can be subsumed under such a large subclass of verbs that we can no longer take responsibility of subclasses specificity. In fact, these two cases are the most difficult to handle and have led linguists to draw different conclusions, e.g. the assumption of different degrees of coextension. This, however, means that the distinction of syntagmas into complements and adjuncts in a sentence (Somers 1984) has almost been abandoned.

It must be stated here that a) - d) is not a complete list of problematic cases. This point needs to be clarified and thus the modified addibility test needs more modification.

Up to now the "modified addibility test" consists in the following:
- If there is free addibility, the syntagma is an adjunct.
- If there are addibility restrictions from other sources than the verb (see a) - c) above) the syntagma is an adjectival.
- All other syntagmas are complements.

1.3 Complement classes

1.3.1 Verbs

The above proposed tests (elimination test and modified addibility test) lead to ten complement classes for German verbs, which can be subclassified further according to their morphosyntactic realization.
There are five case-determined C-classes, including
the prepositional complement. The following table
shows the first 5 C-classes as an example:

<table>
<thead>
<tr>
<th>C-class</th>
<th>syntactic realization</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>co</td>
<td>NP in nominative</td>
<td>Fritz schläft (Fritz sleeps)</td>
</tr>
<tr>
<td></td>
<td>that-clause</td>
<td>Du da hemelt, gedält mir (that you come please me)</td>
</tr>
<tr>
<td></td>
<td>wh-clause</td>
<td>Wer geht, sucht einen Fleiter (Those who leave make a mistake)</td>
</tr>
<tr>
<td></td>
<td>infinitival construction</td>
<td>Promotet zu wokde, let ein schlechtliches lolchte (to be killed in an evil fate)</td>
</tr>
<tr>
<td></td>
<td>main clause</td>
<td>Er sagt, dass ihm fauch (He says this is wrong)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ca</td>
<td>NP in accusative</td>
<td>Er schläft ihn (He sleeps him)</td>
</tr>
<tr>
<td></td>
<td>that-clause</td>
<td>Er schlief, da du kumest (He sleeps, you come)</td>
</tr>
<tr>
<td></td>
<td>wh-clause</td>
<td>Er schlief, wenn du eingekommen bist (He sleeps when you arrived)</td>
</tr>
<tr>
<td></td>
<td>infinitival construction</td>
<td>Er schematic, einen an lassen (He lefts him too enough)</td>
</tr>
<tr>
<td></td>
<td>main clause</td>
<td>Er sagt, dass ihm fauch (He says this is wrong)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ca</td>
<td>NP in genitive</td>
<td>Er schläft dich den (He sleeps the)</td>
</tr>
<tr>
<td></td>
<td>that-clause</td>
<td>Er schlief dich an, dass (He sleeps you too, that)</td>
</tr>
<tr>
<td></td>
<td>wh-clause</td>
<td>Er schlief dich an, dass (He sleeps you too, that)</td>
</tr>
<tr>
<td></td>
<td>infinitival construction</td>
<td>Er schematic, dich an lassen (He lefts you to enough)</td>
</tr>
<tr>
<td></td>
<td>main clause</td>
<td>Er sagt, dass ihm fauch (He says this is wrong)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ca</td>
<td>NP in objective</td>
<td>Ich helfe ihm (I help him)</td>
</tr>
<tr>
<td></td>
<td>wh-clause</td>
<td>Ich helfe, wenn ich will (I help when I want to)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ca</td>
<td>NP in objective</td>
<td>Ich wende mich ihm (I wait for him)</td>
</tr>
<tr>
<td></td>
<td>that-clause</td>
<td>Ich werde dir, dass ich komme (I wait for you to come)</td>
</tr>
<tr>
<td></td>
<td>wh-clause</td>
<td>Ich werde dir, dass ich komme (I wait for you to come)</td>
</tr>
<tr>
<td></td>
<td>infinitival construction</td>
<td>Ich werde dir, dass ich komme (I wait for you to come)</td>
</tr>
<tr>
<td></td>
<td>main clause</td>
<td>Ich werde dir, dass ich komme (I wait for you to come)</td>
</tr>
</tbody>
</table>

There is an sixth class the adverbial complement:
It can be realized as PP as in (7)

(7) Ich wende mich in Sandbrücken.
(He lives in Sandbrücken.)

or as adverbial phrase or as wh-clause.

The seventh C-class is the temporal complement as in (8).

(8) Der Vortrag dauert lange.
(The lecture lasts a long time.),

It can be realized as PP, advP and as wh-clause.
The directional complement as in (9) can also
be realized as PP, advP and as wh-clause.

(9) Er fährt nach Hause.
(He drives home.)

The identifying complement as in (10), can be
realized as NP (nominal) as AP (adverbial), AP
or as whP and as advers.

(10) Er hat schon.
(He is already.)

Finally the verbal complement as in (11) can be
realized as infinitival construction, that-clause,
main clause.

(11) Er fährt ihn kommen.
(He lets him come.)

By using the complement classes and the
syntactic realization by morphosyntactic realizations, all
possible patterns for predicate complement classes
for German can be given. There are approximately 150
patterns which can be subclassified by 4 to 5
syntactic realizations for each complement class.
So, each possible sentence in a realization of
these 150 sentence patterns plus an arbitrary
number of adjectives.

1.3.2 Notes

There is a major difference with the determination
of the category of nouns since there are no obliga-
tory complements with nouns. Thus, only the
modified additivity test can be applied. It must
allow for classifying the uncategorized syntactic: in
(13) as complements of them in (14) as adjectives.

(13) Reisemorgen am Heidelberg (memories of
Heidelberg)
   Heidelberg und Freunden (home and friends)
   Speckgang durch den Wald (walk through the
woods)
   Erwartung, dass der Wirt komme (expectation
that waiter will come)
   Brief aus Ägypten (letter from Egypt)

(14) Der Mann mit dem Kind (the man with the
day)
   Würziger Brot (buttered bread)
   Ein Wissenschaftler wie er (a scientist like him)
   Ein Topf aus Stahl (a pot made of steel)

The modified additivity test rules out the
complements in (13). There are restrictions in
addibility as shown in (15).

(15) *die Mutter auf Frieden (*the mother for peace)
    *der Tisch durch den Wald (*the table through
    the forest)
    *der Tisch, daß der Winter kommt (*the table
    that winter comes)

The adjuncts in (13) are freely addable to all
nouns. There are some semantic restrictions at most
holds for those with verbs.

There are six complement classes for German nouns:
Two case-determined C0 (nominative) and C1 (nps in
genitive, that-clause), prepositional complement,
directional, situative and identifying complement.

1.3.3 Adjectives

The determination of adjective complements works
according to our two tests. There are even some
adjectives which have obligatory complements, e.g.
situated. They are determined by the elimination
test.

There are six complements classes for German
adjectives: four case determined complement classes
accusative, genitive and dative complements and the
prepositional complement, there are quantificatio-
nal complements and the situational complement as
shown in (16):

(16) ziemlich groß (C0)
    (quite large)
    die Arbeit gewohnt (C1)
    (accustomed to work)
    der Arbeit müde (C2)
    (fed up with work)
    den Eltern fremd (C3)
    (alienated from the parents)
    interessiert an Linguistik (C4)
    (interested in linguistics)
    am Fluß gelegen (C5)
    (situated at the river)

1.4 The theoretical nature of dependency

The description of the theoretical nature will
consist of a short description of the relation to
the "higher" level of constituency and the "lower"
level of semantic representation.

1.4.1 Constituency and dependency

The relation between constituency and dependency is
usually regarded as complementary. Constituent-
grammars (C-grammars) are based on the part/whole
relation and define a hierarchical structure with
the respective higher ranking category defined as
being composed of the lower ranking ones. In
contrast to this, the dependency-grammars (D-gram-
mars) define relations between categories of the
same rank, i.e. there is no hierarchical structure
in this ranking sense.

However, this is not sufficient for a description
of the relation between constituency and dependen-
cy. The relation becomes problematic if the C-
grammeme is a recursive subject/predicate grammar
generating a deep tree. In this case, a translation
in the sense of a stratificational M2-system is
very problematic without a tree-flattening
procedure. This procedure could be justified
linguistically because a C-grammar generating flat
trees can generate the same sentences as a C-
grammeme which generates deep trees.

1.4.1 Dependency and semantic representation

The complement/adjunct distinction which has been
made on the syntactic level using purely syntactic
tests is of sentence-semantic importance. Each
semantic representation, be it based on symbolic
logic e.g. Montague Grammar or on a case grammar,
is usually implicitly based on the complement
adjunct distinction. There are, however,
exceptions. E.g. Fillmore's "instrumental" is an
adjunct according to the above mentioned tests
(however, marked as doubtful) since it has the very
same status as the other roles in Fillmore's fram-
work.

The differentiation between complement and adjunct
is made in Systemic Functional Grammar (Halliday,
Fawcett) by distinguishing between participant
roles and circumstantial roles. The participant
role is a semantic interpretation of the comple-
ment-verb relation, whereas the circum-stantial
role is the semantic interpretation of the adjunct-
verb or adjunct-clause relation.

As has been shown (Projektgruppe Verbvalenz 1981)
the complement verb relation can be interpreted
semantically in a lambda categorical grammar. A
semantic description of an expression of natural
language in a lambda categorical grammar consists
of a translation into an expression of the
artificial language lambda L and of a model
theoretic interpretation of this lambda L
expression.

Valency frames of verbs are represented in a lambda
categorical grammar as the number of lambda prefixes
the translated expression receives by the transla-
tion into lambda L.

The lambda operator can bind a variable in its
scope and makes predicates out of sentences.

(17) lambda x [schläft (x)] (ein x sein, das
schläft)
    lambda x [sleeps (x)] (to be an x that
sleeps)

Two-place predicates are represented like (18).

(18) lambda x1 [lambda x2 [betrachtet (x1,x2)]]
    lambda x1 [lambda x2 [looks at (x1,x2)]]

I shall not try to show how such an expression is
interpreted in a model. The point that is made here is that a semantic description in the frame of a lambda categorial grammar uses the syntactic relation between complement and governing verb as the basis of its sentence - semantic description.

2. The role of the syntactic dependency representation level in MT

The facts discussed purely linguistically in section 1 give clear guidelines for application in MT.

2.1 Dependency structure (D-structure) in analysis

The dependency representation (D-representation) serves two purposes:

a) The translation into the D-representation relates the D-structure to the syntagms analyzed on the C-level and thus contributes to the disambiguation of the C-structures which cannot be achieved on the C-level as these can only be reached by the valency statements. (This at least is the case if the two levels are strictly separated.) Usually, the constituent analysis delivers several readings for a longer sentence. Each NP in the genitive case for example, which is an attribute to a noun, has to be interpreted as a potential genitive valency of a verb. In this case, the transition from C-representation to D-representation filters many ambiguous structures by assigning the appropriate d-relation. As there are only a few German verbs with genitive valency, these readings are filtered out in most cases.

b) A second function of the dependency structure is the disambiguation of the verbs (and other elements of word classes which have a valency frame). Different verb readings often are discriminated by different valency frames. An arbitrary example proves this:

(19) anziehen CO
    anziehen CO/Cl
    anziehen CO/Cl/(C3)
    Die Mieten ziehen an.
    (Rents are rising.)
    Fritz zieht die Bremse an.
    (Fritz pulls the brake.)
    Die Mutter zieht dem Kind die Schuhe an.
    (The mother helps the child with the shoes.)

In a sentence in which the verb "anziehen" occurs with only one CO, the reading can be identified unambiguously and translated by t-rules into the IS (Interface) atoms with the corresponding case role. (In EUROTRA, the IS is designed as a semantic interpreted D-structure).

2.2 D-structure in generation

The task of the D-level in generation is the generation of the target language D-structure from IS by assigning the appropriate (correct) surface syntactic valency frames.

In the source language, e.g. German, the verb "sich erinnern" has the syntactic valency frame CO/C4, (which is the complement in the nominative case and a prepositional complement). In the target language English the verb "remember" has the frame CO/C1 (which is the subject and the direct object). In the target language the language-specific surface syntactic valency frame (the direct object) is generated from the interlingual IS. The D-structure is thus a precondition for the generation of correct constituent surface structures according to their valencies.

(20) is an example. (It uses the case roles "processor" and "phenomenon" according to systemic functional grammar.)

The translation undergoes the mentioned levels. Relevant for generation is IS(T) to ERS(T).

(20) ERS (S) #sich erinnern an CO/C4
    IS (S) #sich erinnern an Processor/Phenomenon
    IS (T) #remember Processor/Phenomenon
    ERS (T) #remember CO/Cl

3. The Implementation of D-Representation in the EUROTRA Framework

3.1. Some Remarks on the EUROTRA - Formalism

The most important assumption in EUROTRA is "that translation between natural languages is a sequence of primitive translations between a number of levels of representation" (Arnold et al. 1985b). Such a level of representation is a "language L generated by a grammar G and an interpretation I" (Arnold et al. 1985b). I specifies the syntactically and semantically well-formed expressions of L. G consists of a set of atoms and a set of constructors. Basically atoms are the lexical entries, the constructors are the grammar rules on the different levels. Atoms have a name and a set of features. Constructors have a name and a set of features and a set of arguments which can either be atoms or constructors themselves. They look like (21):

(21) Atom: name {feat1,..., featn}
    Constructor: name {feat1,..., featn} [arg1,... ,argl]
places of the constructor arguments. The adjacent representational levels are related by a translator which is a set of t-rules. It has been said that translations between representational levels are primitive which means that they are (a) compositional and (b) "one shot" (Arnold et al. 1985b).

(a) means that atom is translated to atom and constructor to constructor, where at least the relaxation is allowed that the number and order of constructors differ.
(b) means that the translator takes only well-formed expressions of Gi and yields only well-formed expressions of Gi+1 (Arnold et al. 1985b). This means that there is no internal strategy in the t-rule.

3.2. Some ideas for the implementation of D-structure in the Eurotra formalism

The implementation of D-structure consists of what has been called patterns, in 1 and it has to be done in the formalism sketched in 3.1. That means that for each of the patterns there has to be a set of constructors.

As implementation has just begun the ideas here are somewhat vague, and the proposals certainly do not use all the possibilities the formalism provides.

3.2.1 Sentence Patterns

For sentences the general pattern looks like (22)

\[
(22) S( FD ) \{ \text{gov( FD ), compl1( FD ), \ldots, compln( FD ), adjunct( FD )}* \}
\]

This is just what has been stated above. A sentence S with a feature description FD consists of the pattern with a governing verb with a feature description FD and a complement configuration each complement with a feature description FD plus an arbitrary number of adjuncts each adjunct with a feature description FD.

The above proposed subclassification by syntactic realization can be handled by a cat feature. For pattern C 0/C 1 (sentences with transitive verbs) (23) is an example.

\[
(23) S( FD ) \{ \text{gov( cat=v, \ldots, ), CO( cat=sp, case=nom, \ldots, ), CL( cat=vp, case=acc, \ldots, ), adjunct( FD )}* \}
\]

So, the implementation of the 150 subclassified sentence patterns consists in an enumeration of the S - constructors according to (23).

3.2.2 NP-Patterns and AP-Patterns

For NPs and APs the general patterns look like (24):

\[
(24) NP( FD ) \{ \text{gov( FD ), compl1( FD ), \ldots, compln( FD ), adjunct( FD )}* \}
\]

The subclassification according to syntactic realization has to use the cat feature as in 3.2.1.

4. Conclusion

The D-structure is of major importance for an MT-system and a careful linguistic definition of this level should be made. It is important particularly in a multilingual MT-system like EUROTRA as it is a precondition for IS and thus for transfer. The way in which the D-structure has been presented here, it represents an interface between language specific and interlingual levels. It is interlingual in the sense that the complement/adjunct distinction is regarded interlinguially, and language specific insofar as the classification of complements is language specific.

5. References:


(4) Engel, Ulrich: Syntax der deutschen Gegenwartssprache, Berlin 1982, (Grundlagen der Kasusmorphologie 22)

(5) Engel, Ulrich, Schumacher, Holm: Kleines Va/enzlexikon deutscher Verben, Tübingen 1978, (Forschungsberichte des Instituts für deutsche Sprache 31)

(6) Projektgruppe Verbvalenz: Konzeption eines Worterbuchs deutscher Verben, Tübingen 1981, (Forschungsberichte des Instituts für Deutsche Sprache 45)