Valency Theory in a Stratificational MT-System

Paul Schmidt

IAI Eurotra-D
Martin-Luther-Str. 14
6600 Saarbrücken
West Germany

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
and
(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. LFG 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 syntagma.

Those valency-bound syntagmas are the complements. The syntagmas which are not valency-bound are the adjuncts. All syntagmas 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 syntagmas the wordsubclass specific elements are valency-bound (Engel 1982). Wordsubclass specific elements are those syntagmas which can occur only with elements of wordsub-classes.

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 adaption 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 syntagma 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 interlingually, which is not unimportant for EUROTRA.

The "free addibility test" consists of checking which syntagmas can 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 explanations 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 adjunct classes with certain subjects.

(2) *Der Baum steht absichtlich am Hang.
(*The tree is 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 form correct sentences containing both the verb and the adjunct. Therefore, there is no principled impossibility of their combination.

(3) Er kommt 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 adjunct classes are in fact not combinable with a verb. In

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

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

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

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

d) However, there are certain problematic cases.

(6a) *Es regnet mit einem Hammer.
(*It is raining with a hammer.)

(6b) *Das Auto ist rot bis morgen.
(*The car is red until tomorrow.)

Instrumental adjuncts and certain temporal adjuncts can scarcely be described as being freely addable as they only combine with action verbs and so fulfil 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 talk reasonably of subclass specificity. In fact, these two cases are the most difficult to handle and have led linguists to draw different conclusions, as e.g. the assumption of different degrees of cohesion. This, however, means that the distinction of syntagmas into complements and adjuncts in a sentence (Somers 1984) has simply been abandoned.

It must be stated here that a) - d) is not a complete list of problematic cases. This point needs to have more investigation 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 restriction from other sources than the verb (see a) - c) above) the syntagma is an adjunct.
- 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:

C-class	syntactic realization	example
C0	- NP in nominative	Fritz schläft (Fritz sleeps)
	- that-clause	Daß du kommst, gefällt mir (That you come pleases me)
	- wh-clause	Wer geht, macht einen Fehler (Those who leave make a mistake)
	- infinitival construction	Ermordet zu werden, ist ein schreckliches Schicksal (To be killed is an awful fate)
C1	- NP in accusative	Er schlägt ihn (He beats him)
	- that-clause	Er sieht, daß sie kommt (He sees that she is coming)
	- wh-clause	Er sieht, wen du eingeladen hast. (He sees whom you invited)
	- infinitival construction	Er glaubt, genug zu haben (He believes he has enough)
	- main clause	Er sagt, das ist falsch (He says this is wrong)
C2	- NP in genitive	Er erinnert sich des Fehlers (He remembers the mistake)
	- that-clause	Er klagt ihn an, daß er einen Mord begangen hat (He accuses him of having committed murder)
	- wh-clause	Er beschuldigt ihn, wessen er sich erinnert (He accuses him of that which remembers)
	- inf.constr.	Er beschuldigt ihn, ein Mord begangen zu haben (He accuses him of having committed murder)
C3	- NP in dative	Ich helfe ihm (I help him)
	- wh-clause	Ich helfe, wem ich will (I help whom I want to)
C4	- PP	Ich warte auf ihn (I wait for him)
	- that-clause	Ich warte darauf, daß er kommt (I wait for him to come)
	- inf.constr.	Ich warte darauf, eine Chance zu erhalten (I wait for (getting) a chance)

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

(7) Er wohnt in Saarbrücken.
(He lives in Saarbrücken.),

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 (nominative) as NP (accusative), AP or as wie-P and as als-P.

(10) Er ist doof.
(He is stupid.),

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

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

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

1.3.2 Nouns

There is a major difficulty with the determination of the valency of nouns since there are no obligatory complements with nouns. Thus, only the modified addibility test can be applied. It must allow for classifying the underscored syntagmes in (13) as complements but those in (14) as adjuncts.

(13) Erinnerungen an Heidelberg (memories of Heidelberg)
Hoffnung auf Frieden (hope for peace)
Spaziergang durch den Wald (walk through the forest)
Erwartung, daß der Winter kommt (expectation that winter will come)
Brief aus Ägypten (letter from Egypt)

(14) Der Mann mit dem Hund (the man with the dog)
Peters Hand (Peter's hand)
Ein Wissenschaftler wie er (a scientist like him)
Ein Topf aus Stahl (a pot made of steel)

The modified addibility 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 genitve, 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 complementclasses for German adjectives: four case determined complement classes accusative, genitive and dative complements and the prepositional complement, there are quantificational 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-grammars) 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 dependency. The relation becomes problematic if the C-grammar is a recursive subject/predicate grammar generating a deep tree. In this case, a translation in the sense of a stratificational MT-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-grammar 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 framework.

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 complement-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 categorial grammar. A semantic description of an expression of natural language in a lambda categorial grammar consisits 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 categorial grammar as the number of lambda prefixes the translated expression receives by the translation 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 syntagmas 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 C0	Die Mieten ziehen an. (Rents are rising.)
	anziehen C0/C1	Fritz zieht die Bremse an. (Fritz pulls the brake.)
	anziehen C0/C1/(C3)	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 C0, the reading can be identified unambiguously and translated by t-rules into the IS (Interface) atom 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 assingning the appropriate (correct) surface syntactic valency frames:
In the source language, e.g. German, the verb "sich erinnern" has the syntactic valency frame C0/C4, (which is the complement in the nominative case and a prepositional complement). In the target language English the verb "remember" has the frame C0/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 C0/C4
	IS (S)	#sich erinnern an Processor/Phenomenon
	IS (T)	#remember Processor/Phenomenon
	ERS (T)	#remember C0/C1

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,...,argi]

A constructor is syntactically well-formed if its arity equals the number of its arguments and if the arguments are well-formed. It is semantically well-formed if its arguments unify with the argument

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 } [gov{ FD }, compl1{ FD },...,
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 } [gov {cat=v,....}, C0 {cat=np,
case=nom,....}, C1 {cat=np,
case=acc,...},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 } [gov { FD }, compl1 { FD },...,
compln { FD },adjunct{ FD }*]

AP { FD } [gov { FD }, compl1 { FD },...,
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 lingustic definiton 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 interlingually, and language specific insofar as the classification of complements is language specific.

5. References:

(1) Arnold,D., des Tombes,L., Jaspaert,L.,:
Eurotra Linguistic Specifications, (manuscript), 1985, (=1985a)

(2) Arnold,D., et al.: A Mu-1 View of the <C,A>,-T Framework in EUROTRA, in: Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages. Colgate University, Hamilton, New York, August 1985, pp 1 - 14. (=1985b).

(3) Arnold et al.: The Eurotra Reference Manual, (manuscript), February 1986

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

(5) Engel, Ulrich, Schumacher, Helmut: Kleines Valenzlexikon deutscher Verben, Tübingen 1978, (Forschungsberichte des Instituts für deutsche Sprache 31)

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

(7) Somers, Harold: On the Validity of the Complement-Adjunct Distinction in Valency-Grammar, in: Linguistics 22, 1984, pp 507-530