5

1965   International Conference on Computational Linguistics

MACHINE TRANSLATION AND CONNECTEDNESS BETWEEN PHRASES

Karel   Č U L Í K

Mathematical Institute of Czechoslovak Academy

of Sciences, Žitná 25, Praha 1, Czechoslovakia

1. Machine Translation.

To translate a special text from one language into another means to construct to the given text in the first language, such a text in the second-one that has the same meaning (the same is told in it) like the given text.

The translation from a natural language L_1 into a language L_2 is such a function F (in the more general case it is a many-valued function) which assigns to any text T in the language L_1 such a text F(T) in the language L_2, that F(T) has the same meaning like T. If we introduce a semantics (or interpretation) M as a function (in general also many-valued) which to any expression E of some language L assigns its meaning $M[E]$ (compare [4]), it is possible to say that the function F assigning the texts F(T) from L_2 to the texts T from L_1 is the translation only when

$$(1) \qquad M[T] = M[F(T)] \quad \text{for any text} \quad T \quad \text{from } L_1.$$

At the machine (or automatic) translation, the matter is, to define the function F as a mechanizable procedure (i.e. algorithm) according to which an arbitrary starting text T in L is being succedingly modified till we get the translated text F(T) fulfilling naturally (1). The corresponding algorithm can be finally programmed for a suitable computer. With respect to the used computer, the programme of the algorithm must not be too long, not even the wide range of memories must not be emploied and at last the translation must not take up too much of time. Usually it is required for the algorithm of a translation, to be the most effective.

2. Translation "Sentence by sentence".

The translation F is theoretically - as every function - defined as a set of pairs $[T, F(T)]$ where T runns through all texts in L_1. If we really had these pairs practically at disposal, we could use a trivial algorithm of the translation F: we should put in the memory of the computer all the pairs $[T, F(T)]$ and when being given the starting text T, we should find out in the memory the pair, in the first place of which T would be situated, thus the pair $[T, F(T)]$, and the demanded translation F(T) would be on the second place of this pair. This is, of course, not only funny but also impossible.

It seems to be funny because of the fact that to have practically at disposal the pairs $[T, F(T)]$ it would mean to use lively translation and thus to translate all possible texts in advance. But the automation of translation signifies to exclude as much as possible the direct intervention of man out of the proceeding of the translation and thus to sustitute a man by a machine. On that score, we do not possess practically the pairs $[T, F(T)]$.

It seems to be impossible because the texts T are too many (it would be possible to admit that infinitely many) and the pairs $[T, F(T)]$ could not be included in any computer. On the other hand, it is necessary to admit that it concerus the algorithm, which is very simple (only to look up in the memory would take up too much of time).

The trivial algorithm being practically impossible, it is necessary to try to decompose long texts into parts and then to translata part by part. Naturally, it seems to be profitable, to treat sentences, that are in printed texts distinctly separated by points, as these parts. Thus, every text T is a sequence

of partial texts, i.e. sentences S_1, S_2,...,S_k so that we may write that $T = [S_1, S_2, ..., S_k)]$. The function of the translation F is, of course, according to the asumption defined for all texts and thereby also for particular sentences S_1, S_2,...,S_k so that it is possible to construct a somposed text $[F(S_1)$, $F(S_2),...,F(S_k)]$ from the translations of these sentences $F(S_1)$, $F(S_2)$, ..., $F(S_k)$, that are some partial texts in L_2. At the same time, the translations of sentences follow in the same sequence as did the starting sentences in the text T. It may happen - and we should sure welcome it, if it were always - that it holds

(2) $\qquad F(S_1 \circ S_2 ... S_k) = [F(S_1) \circ F(S_2) \circ ... F(S_k)]$

or at least the weaker condition

(3) $\qquad M[F(S_1 \circ S_2 ... \circ S_k)] = M[(F(S_1) \circ F(S_2) ... \circ M(S_k)]$.

From (2) there follows (3) but in no way the contrary. For the translation the condition (3) is sufficient. It might, namely, happen that we translate the text T, as a whole, differently than when translating it succeedingly in parts $S_1, S_2, ..., S_k$, so that (2) does not hold, but despite this (3) holds.

In the condition (2) and similarly the condition (3) were fulfilled for any text $T = (S_1 \circ S_2 ... S_k)$ in L_1, it would signify that it was always possible to translate single sentences of the text quite independently each of another. It is probably not true. Sometimes, it is necessary to know, how the sentence S_1 was translated, when we want to translate correctly the sentence S_2, becau-

se regularly both sentences are connected as to the contents, and not always this connection is expressed by syntactical means. In addition, sometimes it is necessary to translate too long sentences from L_1 as two or more sentences from L_2 and then not even the sequence is possible to be defined in advance.

In spite of this, the condition (2) or at least (3) is the basic asumption for any translation "<u>sentence by sentence</u>" and most part of translation belongs to such a type of contemporary translations. To be competent to accept the asumption (2) or (3) it sufficies to confine oneself to some texts only, and the texts not fulfilling this assumption are necessary to be adapted before the translation in order to make them able to fulfill it. It is not clear, of course, how to find it out at the given text, before starting the translation.

The asumption (3) stands for nothing else than

(4) $M(T) = M\left[F(S_1) \cdot F(S_2) \ldots F(S_k)\right]$ for any text $T = \left[S_1 \cdot S_2 \ldots S_k\right]$

and hence the substantial simplifying of the definition of translation can to be seen. It suffices, namely, to suppose that it is necessary to define only a partial function F^* of the function F, that is defined for arbitrary sentences in L_1 only (on no account for arbitrary texts when the sentences are a special case of simple texts). Thus, there holds $F^*(S) = F(S)$ for every sentence from L_1 and out of (2) there follows

(5)
$$F(S_1 \cdot S_2 \ldots S_k) = \left[F^*(S_1) \cdot F^*(S_2) \ldots F^*(S_2)\right]$$

so that in fact we are able to cope with the function F^* when trans-
lating the texts.

Similarly, like at the function F, it is possible at the
function F^* too, to try to give a trivial algorithm making
full use of all pairs $\left[S, F^*(S) \right]$. But the situation improves
only a little. It wold be again necessary to translate all the
sentences in advance and even these are still too many, so that
all preceding reasons remain valid, what means, that it is necessary
to try to decompose even the sentences in parts, and to trans-
late the sentences in parts, too.

3. Translation "word by word".

By the decomposition of the translation into sentences
there were no difficulties because in printed texs this de-
composition into sentences was just ready, and according to
the syntanctical means it was possible to decompose its parts.

Of the same simplicity and uniqueness is the decomposition
of the sentence S in its single words W_1, $W_2, \ldots, W_k$ separated
by interspaces, so that it is possible to write S = $(W_1 W_2 \ldots W_k)$
like at the text.

Besides, the linguists have constructed, a long time ago,
a binary translative dictionary from the language L_1 into the
language L_2. This dictionary is, in fact, defined as a set of
the pair of words, the first-one from L_1 and the second-one
from L_2 having the same meaning. If we denote by f the trans-
lation from L_1 into L_2, where f is again generally many-valued

function (thanks to the homonymy of words) f evidently fulfils the condition (in details see $[\ 3\ ]$)

(6) $\qquad M[f(W)] = M(W)$ for every W from L_1 .

In whatever way the trivial algorithm of the translation of texts and sentences was funny and impossible, this algorithm is in case of the translation of words not only possible, but also it is used from time to time by living translators. At most machine translations, there is really choosen for the algorithm of the translation f the just mentioned trivial algorithm, i.e. into the memory of a computer there are input all pairs W, f(W) and the most tedious procedure is - how it was said- to look up in the memory and to compare.

But it is not necessary to use this trivial algorithm. It is possible to construct a sequential automaton and thus to construct also a corresponding technical apparatus which will realize the function f, i.e. if there enters on its input the word W as a sequence of letters $l_1 l_2 \ldots l_k$, which is eventually prolongated by the means of several help-symbols (compare $[\ 5\ ]$) we get on output again the sequence of letters $l_1^* l_2^* \ldots l_k^*$, that eventually starts with several help-symbols and simultaneously there holds that

(7) $f(l_1 l_2 \ldots l_j) = 1_i^* 1_i^* \ldots 1_k^*$, when $W = l_1 l_2 \ldots l_j$, $j \leq k$.

There is a question, whether there is not possible, when using this automaton, to shorten the time necessary for translating, when evidently all lost times can be excluded at looking up in the input dictionary.

Taking no account to the fact, in what way the translation of words f is given, one may ask, whether for the sentence $S = (W_1 W_2 ... W_k)$ there holds

(8) $F^*(S) = \left[f(W_1)\ f(W_2)\ f(W_2) \circ \circ \circ f(W_k) \right]$,

which is the similar condition to the condition (2) for the translation of texts.

It is known that this condition holds nearly never for most natural languages, because the translation of words f is the translation of words in basic form only, whereas by the decomposition of the sentence $W_1, W_2, ... W_k$ are in various word-forms. To put it differently:the function f respects only the lexical meaning of words but does not take into consideration morphological questions.

However even in the case, the function f could be prolongated from basic forms on other form of words-what, of course, need not be possible-or on the contrary, if we adapted to the basic form single words W_i in the decomposition of the sentence, yet despite this all - even under these suppositions - the condition (8) would be fulfilled in the case only, that there are concerned two languages L_1 and L_2 that are very strongly cognate, or two dialects of the same language, or in the case L_1 and L_2 are not cognate, the considered sentence S must be very simple.

The translation fulfilling the condition analogous to the condition (8) may be called the translation "word by word". But unfortunately it is known that such a translation is im-

possible in natural languages, although it would be very advantageous and simple.

It does not mean, of course, that the decomposition of the sentence into words cannot be used; it is too fine and therefore it is necessary to decompose the sentence in another way, at all events in such a way that single parts will contain more than one word (and we suppose, under tacit consent, that the words have full meaning, not only being help-words with the grammer meaning) and that these parts need not be sentences. Before introducing these parts, it is necessary to take into consideration various necessary morphological and grammatical statements, and thereby to adapt properly the condition (8), too, where there were no differences between the basic word-form or the mere stem of the word and its various possible forms.

4. Syntactical and semantical characteristics of words.

First of all we may suppose that, to every word-form (shape) which appeared in the decomposition of some sentence into words, we are able to define its <u>basic form</u> or <u>stem</u> W (the function f refers just to these basic forms) and its <u>characteristic</u> $c = (x^1, x^2, \ldots, x^n)$, where n is according to the need a sufficiently great integer and single x^j are some grammatical morphological and eventually even other data referring to the form W. For example x^1 can be the datum on word-kind, x^2 the datum on case, x^3 on gender, x^4 on number, x^5 on person, x^6 on time, x^7 on mood a.s.o.

Naturally we also assume that on the contrary, if it is given the basic form w and prescribed (of course admissible) the characteristic c, it is easy to define the starting form W. Thus, we suppose that there are given functions h and h^{-1} such that $h(W) = (w,c)$ and $h^{-1}(w,c) = W$ in each of considered languages (in L_1 and L_2 it will be functions h_1 and h_2).

Although it is well known that the sentence and its translation need not have the same number of words, this demand is not far from truth when we pay attention only to the full-meaning words. Let us consider, however, such sentences S which fulfil this demand (this is the supposition for making the comment easy) i.e. if it is $S = (W_1 \ W_2 \ \ldots \ W_K)$ then

$$(9) \qquad F^*(W_1 \ W_2 \ \ldots \ W_K) = (\overline{W}_1 \ \overline{W}_2 \ \ldots \ \overline{W}_K) \, ,$$

where evidently $\overline{W}_1$ are the words of the decomposition of the sentence $F^*(S)$. The task is, how to get the chain $(\overline{W}_1 \ \overline{W}_2 \ \ldots \ \overline{W}_K)$ from the given chain $(W_1 \ W_2 \ \ldots \ W_K)$ and we know already that it is not possible to get it by means of the translation word by word according to (8) directly.

When we use a function h_1 we get from any word-form W_i its basic form w_i and characteristic c_i, namely, $h_1(W_i) =$ $= (w_i, \ c_i)$ and thus we can differentiate the data $(W_1, W_2, \ldots, W_K)$ and data $(c_1, c_2, \ldots, c_K)$.

It is similar with the translated sentence $(\overline{W}_1 \ \overline{W}_2 \ \ldots \ \overline{W}_K)$ when one uses the function h_2. Again we are able to discern data $(\overline{W}_1, \overline{W}_2, \ldots, \overline{W}_K)$ and data $(\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_K)$ when evidently $h_2 (\overline{W}_i) = (\overline{w}_i, \overline{c}_i)$.

Now, it is clear that instead of the condition (8) ought to be the condition

$$(10) \qquad \overline{w}_i = f(w_i) \qquad \text{for} \quad i = 1, 2, \ldots, k$$

because here f is really used for basic forms of words. Then the translation fulfilling (10) is in fact the translation „word by word" but only in the respect of the meaning of word, being far from complete translation. There is missing the proceeding, how to get from the starting characteristic $(c_1, c_2, \ldots, c_K)$ characteristic $(\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_K)$. And just here there is impossible – except for the most simple example – to find such a function g in order to hold

$$(11) \qquad \overline{c}_i = g(c_i) \qquad \text{for} \quad i = 1, 2, \ldots, k.$$

If it were the case, or in these cases for which the function g could be found, the translation would be easy, because it would evidently hold

$$(12) \qquad \overline{W}_i = h_2^{-1} \left[f(w_i), \; g(c_i) \right] \text{ for } i = 1,2,\ldots, k,$$

where, of course, $h_1(W_1) = (w_i, c_i)$ for $i = 1,2,\ldots, k.$

As it is impossible to translate one characteristic after another, it is necessary to use instead of the supposed (but in general not existing) function g, the more complicated function G. This function will not define single characteristics $\overline{c}_i$ in dependence on the sole characteristic c_i as it was to be in (11), but in dependence on all characteristics $(c_1, c_2, \ldots, c_K)$ so that it may be written analogically to (9)

$$(13) \qquad G(c_1, c_2, \ldots, c_K) = (\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_K),$$

where, properly, would be necessary to differentiate functions G_1, G_2, G_K and put $\overline{c}_i = G_i (c_1, c_2, \ldots, c_K)$ for $i = 1,2,\ldots k.$

Whereas the condition (10) has been fulfilled quite frequently (especially at simple sentences and above all when we weaken it by admitting the changed the ordering of words $\overline{W}_i$ in comparison with words W_i), i.e. it is often possible to translate word by word as for the meanings of single words, the condition (11) has nearly never been fulfilled. It can be understood, because in the respect of meaning the languages do not differ as a matter of fact and this matters in (10), while morphologically and eventually even grammatically single languages differ very strongly and these facts matters in (11).

From this also follows that difficulties of impossibility of the translation word by word according to (8) are - for the differentiation of the meaning and characteristic - due to the characteristic and not to the meanings. This fact has a considerable heuristic import. It is, namely, evident that the suitable parts into which we want to decompose the sentences are to be found with respect to their significance and not with respect to their syntactical or even morphological properties. In this case, namely, these parts will be found at the same time in all languages even if having been expressed in different languages by different syntactical an morphological means. It is naturally self-evident that between the significance of considered parts and their syntactical expressions are close connections (see [2]).

Because of this, it is necessary to introduce, besides the mentioned characteristics some others more, namely, logical and semantical that will be common for all languages and will be quite independent of the syntax of languages. And just this condition is fulfilled by the logical and semantical questions.

Under the logical characteristics of words we understand data on the fact whether and what logical conjuctions or other logical means (as quantors or negations) are by these words expressed. These facts are known from the logical analysis of sentences worked out by R. Carnap.

By the senmantical characteristics of words we understand data on the fact whether the given word (we suppose word with full meaning, in no way grammatical or logical words) plays the role of individual constant, or variable i.e. whether it defines a certain object

(here the term object is used in the wide sense of the term) or
an arbitrary-one with certain propertied, or plays the role of
one-placed predicat, i.e. denotes some property, or of two-placed
predicat, i.e. denotes two-member relation, or in general n-placed
predicat, i.e. denotes n-figured relation.

The situation is not so hopelessly complicated as it would
seem at the first sight. For instance, the individual constants
even the variables are only substantives while verbs are always
predicats one-. two- three- even more placed, according to the
smaller or greater number of their objects. Adjectives are always
one-placed predicate a.s.o.

Besides the mentioned - and in the logic current - it is
necessary to consider as semantical characteristics data on time
and place and probably not yet quite distinctly defined data
referring to the conditions under which the situation is being
described (here belong some adverbial modifier).

Thus, we suppose that we know the function k (analogously
like h) which to any word-form W assigns its <u>logical</u> and
<u>semantical characteristic</u> d, thus, $k(W) = d$, while again
$d = (y^1, y^2, ..., y^K)$ where e.g. y^1 is a datum, whether the
word W is the logical functor and of what kind, y^2 is a datum
whether the word W is predicate and how many-placed, y^3
a datum of what kind the predicate is, y^4 whether there is in W
the definition of time and of what kind, y^5 whether W is the
definition of place and of what type, y^6 a datum on the special
definition of the condition a.s.o.

5. Primitive phrases.

With respect to the semantical characteristics of words it is possible and quite natural - similarly as in the predicate-logic - to take as important in a sentence such phrases (i.e. their parts) that include always one word playing the role of n-placed predicate - so called <u>basic predicate</u> - in this phrase (while phrases contain n+1 words) and the other words (just in the number n) play the role of individual constants or variables being placed on single places (positions) of the considered predicate. In accordance with the mentioned it referrs, in this phrase, to the denotation of the n-membered relation and to the denotation of all n objects that are mutual in this relation. Thus, every such phrase is, in fact, a certain statement or a definition on the situation. As it is the analogy of the primitive formula in the mathematical logic (e.g. $P(a_1, a_2, a_3)$ is the primitive formula when we know that P is a three-figured predicate, that x_1, x_2, x_3 are individual constants placed on their three places, and that the mentioned record says that the objects denoted by these constants are in relation denoted by the predicate P) we call such a phrase the primitive phrase.

Simultaneously, it is immediately evident that the primitive phrase need not have the grammatical form of a sentence, and in most cases it really does not have it. It has the form of a sentence just when its predicate is a verb and when this verb has not the grammatical form of a gerodnd or a participle. For instance „a man reads a book" is the primitive phrase in the form of a grammatical sentence (here evidently „reads" plays the role of two-figured predicate), but the primitive phrase „a man reading book" or „a man who

is reading a book" has not the form of a sentence even if having the same meaning like the preceeding phrase, because both express the same fact. Some other types of primitive phrases are e.g. these „very good" where „very" is a one-placed predicate and on the place of it stands „good" (although in another primitive phrase „good book" is good itself a one-placed predicate), or „reads quickly", where a one-placed predicate is quickly a.s.o.

From these examples there follows that primitive phrases correspond with primitive formulas in the predicate logic of higher order. In the phrase „man is mortal" there is evidently concealed the universal quantor „every" so that this phrase has the same meaning like „every man is mortal" and thereby to not a primitive phrase but a composed-one.

With regard to the syntactical side, the primitive phrases differentiate on the basis of the characteristics of single words. For instance, the sequence of word-characteristics (c_1, c_2, c_3) where c_1, c_2, c_3 are such that c_1 denotes a substantive in the first case, c_2 denotes a transitive verb, c_1 and c_2 simultaneously coinciding in their components as for the gender and number, and at last c_3 signifies that it referrs to a substantive in the fourth case, when, in addition, the range of characteristics sets the future word-order, is the characteristic of the primitive phrase „a man reads a book" (when an indefinite article is a considered not to be self-contained word, and we assign it always to the word suceeding it immediately). If we, namely, made full use of the function h we would get h [a man] = [man, c_M], h [reads] = [read, c_R] and h [a book] = [book, c_B] , and

simultaneously it couls sure hold $c_M = c_1$, $c_R = c_2$ and $c_B = c_3$.

The considered language is, besides its word-store put down in the dictionary, characterized also by the list of all characteristics of primitive phrases. That is what we shall suppose at any of the considered languages.

If we use for the considered primitive phrases. That is what we shall suppose at any of the considered languages.

If we use for the considered primitive phrase „a man reads a book" the function k, we get some semantical word-characteristics d_M, d_R and d_B, and one of them will be especially distinguished as a basic predicate of the considered type of the primitive phrase. In our case it is d_R and for illustration we shall come to an agreement that this basic predicate and other semantical characteristics will be put down in the same way like it is done in the predicate logic, namely $d_R(d_M, d_B)$. Let us call this entry the semantical characteristic of the considered phrase and also the semantical characteristic corresponding with the syntactical characteristic (c_M, c_R, c_B) of the considered phrase. Analogous agreements are to be made even with respect to other components of the semantical characteristic, particularly for the definition of time, place and other conditions.

At the same time, the semantical characteristic comprises these facts: d_R a datum that it referrs to the two-figured predicate (eventually specialized by the denotation of some activity), on the first place of which is just the word, with the semantical characteristic d_M (e.g. with a supplement „agens") and

the second place of which corresponds with the semantical characteristic d_B (eventually with supplement „patiens"). Simultaneously, the semantical characteristics may be eventually complemented with further data, when it turns out to be suitable. It is important only, that there are to be data (as it is mentioned in the parentheses) that referr to the meaning and that are common for all languages.

By the semantival characteristic $d_R(d_M, d_B)$ there is put down, in the basic semantical categories, just what we want to express (By that time the basic forms Read, Man, Book are failing; these are possible to be chosen differently) while the corresponding syntactical characteristic (c_M, c_R, c_B) puts down how to express it.

Now, the way is evident, how to translate primitive phrases from the language L_1 into L_2. There is important that we suppose that whatever can be expressed in L_1, can be expressed even in L_2 what is the basic supposition on the possibility of translating. From this there follows for the function of the translation of words f that for every basic form W from L there exists $f(W)$ in L_2, and that for every semantical characteristic $d_1(d_2, d_3, \ldots, d_n)$ corresponding with the syntactical characteristic $(c_{i_1}, c_{i_2}, \ldots, c_{i_n})$ in L_1 thete exists the syntactical characteristic, corresponding with it $(c_{j_1}, c_{j_2}, \ldots, c_{j_n})$ in L_2 and just this-one (of course they may be several) will be declared to be the trabslation of the corresponding syntactical characteristic from L_1. Thereby a further function φ is defined (eventually a many-valued) for which there holds that

(14) $\quad \varphi \, (c_{i_1}, \, c_{i_2}, \, ..., \, c_{i_n})^1 = (c_{j_1}, \, c_{j_2}, \, ..., \, c_{j_n})^2 \, ,$

when $(.....)^1, \, (.....)^2$ have the same semantical characteristic.

Now, if it is given a primitive phrase $(W_1, \, W_2, ..., \, W_n)^1$ in L_1 we use first the function h and we get $h \, [(W_i)] =$

$= [w_i, \, c_i]$ for $i = 1, 2, ..., \, n,$ and thus its syntactical

characteristic $(c_1, \, c_2, ..., \, c_n)^1$ but now $\varphi \, (c_1, \, c_2, ..., \, c_n)^1 =$

$= (c_{j_1}, \, c_{j_2}, \, ..., \, c_{j_n})^2$ so that

(15) $\quad F^*(W_1 W_2 ... \, W_n) = (h_2^{-1}[f \, (w_{j_1}), \, c_{j_1}] \, h_2^{-1} \, [f(w_{j_2}), \, c_{j_2}] \, ...$

$$... \, h_2^{-1} \, [f(w_{j_n}), \, c_{j_n}] \,)$$

and this is, in fact the needed weakening of the condition (12).

It is evident that the primitive phrase from L_2 on the righthand side of the equation (15) has really the same meaning as the primitive phrase on the left-hand of the equation. As for the meaning of single words, this is guaranteed by the function f and as for the meaning of the whole phrase, it is guaranteed by the function φ, that fulfills (14) that here in a special case plays the role of the function G, because (14) and (13) are identical.

6. Compound phrases.

The composing of two or more primitive phrases in compound phrases can be performed by the usual logical means (e.g. by means of logical conjunctions "even" or "if,...then..." eventually of other analogous conjunctions like "but" and similar, or by means of negation and quantors) in a well-known way, or can be perfomed by the pure linguistic expressive means. Both these kinds may be arbitrarily interchanged by the successive composing.

Similtaneously, it is decisive that the composing of single primitive phrases corresponds with the composing of their syntactical and, of course, semantical characteristics. Besides, there are mostly composed such two phrases that have some word in common, or where some word is repeated. This fact is necessary to be distinguished especially by composing the corresponding characteristics, or - what is in substance the same- it is necessary to join to any word-characteristic c the symbol expressing a variable for the basic forms of word, so that we shall write $\ulcorner X, c_1 \urcorner$ where we can put for X the real basic forms of words.

The composing of primitive phrases in compound phrases belong to the field of the synthesis of phrases. If we, for instance, want to say that some man reads a book and simultaneously that he reads quickly and, in addition, that this book is good and even very good, we can express it in the following compund phrases P = (which is grammatically the form of the sentence) "man reads quickly a very good book". The syntax of this phrase is not evidently expressed by the logical means.

In the considered case the following primitive phrases are concerned: P_1 = " a ma reads a book, "P_2 = " reads quickly", P_3 = "good book" and P_4 = "very good". From these phrases the compound phrase is put together. If we use the function h for single words of the considered phrase, we get successively $h[a \ man] = [man, \ c_1]$, $h[quickly] = [quick, \ c_2]$, $h[reads] = [read, \ c_3]$ ...,

,,,,$h[a \ book] = [book, \ c_6]$, where c_1, c_2,..., c_6 are the corresponding syntactical characteristics. Analogously when we use the fun - ction k we get the semantical characteristics of single words.

The syntactical characteristics of single (separate) primitive phrases P_1, P_2, P_3, P_4 are successively C_1, C_2, C_3, C_4, where $C_1 = ([W_1, \ c_1], [W_3, c_3], [W_6, \ c_6]$, $C_2 = [W_2, \ c_2], [W_3, c_3])$ $C_3 = ([W_5, c_5], [W_6, \ c_6])$ and $C_4 = ([W_4, \ c_4], [W_5, \ c_5])$. At the same time, there is very important that some variables W_i occur simultaneously in two primitive phrases. Thereby is, namely, expressed the circumstance that by these two phraes is told so- mething of the same fact and just this circumstance plays the decisive role at stating the constents-connection among more phrases.

If we started from the given phrase " a man quickly reads a very good book" we would find the mentioned four primitive phrases as follows: first we would use for single words the functi- on h and k and then we would find for every word of the mentioned phrase, which can be the basic predicate of some primitive phrase (it can be found out of its semnatical characteristics and by the semantical characteristics of the primitive phraes) further words belonging to it in a certain primitive phrase, i.e. which take places of the considered predicate and this occurs only

in comparison of the syntactical characteristics of words
from the given phrase with the syntactical characteristics
of a certain investigated primitive phrase.

For instance, in our case, if we have found out according
to the semnatical characteristic that the word "reads" is two-
placed predicate, we would find out the syntactical characteris-
tics of such primitive phraes, the basic word of which was
just the two-placed predicate. Then we have known what syntac-
tical characteristics of words and - as far as the word-order
is concerned - where there are to be found, so that we find out
whether the investigated primitive phrases are in the given phra-
se included. When finishing it for all these words, we shall find
it successively for all primitive phrases that in the given
phrase are comprised.

In such a way is, namely, depicted the analysis of the
compound phrase, not composed by the logical means. If there are
used the logical means, then the given phrase is decomposed
like in the logic.

But it is necessary to mention in addition, that for the
economy-reasons and for saving the number of syntactical charac-
teristics of the primitive phrases, it is convenient to work
often with incomplete characteristics only. The question
is, whether we shall include two primitive phrases" a man reads
a book" and "the man read a book" into one (incomplete)
syntactical characteristic, or into two different and naturally
complete -ones. The incompleteness will consist in the failing
fact on number (and similarly it would be in other phraes with
data on gender and case), but naturally there would not fail

the datum on coincidence in number between „a man" and „reads
or „the man" and „read" because this fact will be just decisive
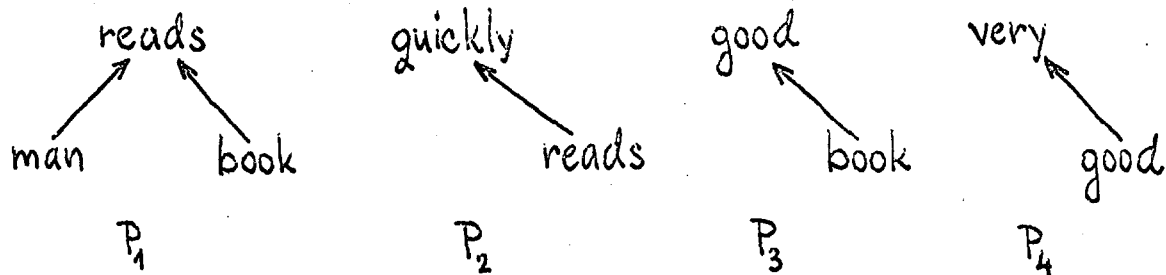for the incomplete characteristics.

The possibility of the use of incomplete syntactical cha-
racteristics by the synthesis is, of course, also evident. If we
want to make the whole synthesis of the compound phrase indepen-
dent on proper meanings of single words, then we can give in the
syntactical characteristics neither the gender nor the number,
because both of them are defined differently no sooner than by
the choice of the basic form (because in many cases genders are
steadily fixed). But even here it is not the matter of principle
but the matter of effectivity.

7. Semantical dependence and connectedness.

As, according to the supposition, there is denoted in
every primitve phrase its basic predicate which always stands in
front of parantheses in its semantical characteristic (for instan-
ce at $Q(x,y)$ Q is the basic predicate) it is possible to define
the semantical dependence among the words of the primitive phrase
by the demand that the basic predicate always depends on all
other words that occur in the phrase, i.e. on its arguments (e.g.
Q depends on x and on y). Just so justified would be the defi-
nition that, on the contrary, all arguments depend on the basic
predicate.

If we demonstrate this semantical dependence on a diagramm,
we always draw the connecting line, provided with an arrow-head,
directing from an argument to a basic predicate. At the same time,

of course, according to the position of separate words - if
they are more to the left or to the right - we discern always
the word-order. Four primitive phrases from the preceeding
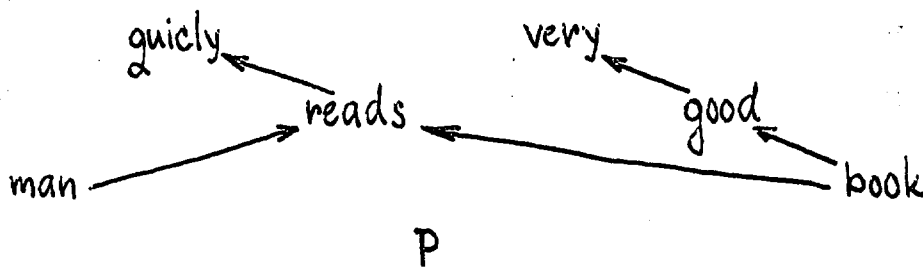paragraph are deomstrated in the following diagramm:

We say that the semantical dependence is concerned because this
relation among the basic predicate and its arguments, expressed
just by the theorem that objects denoted by the arguments are
in relation defined by the basic predicate, is quite initial
definition referring evidently to the reality. The semantical
dependence does not refer to anything else than to the fact of
telling something of something (on the mathematical level the
fact of telling $Q(x,y)$ can be transferred only on the basic
relation of the adherence to the set) when one passes from the
predicate Q to the binary relation Q^* and puts down
$(x,y) \in Q^*)$.

We say further that in the primitive phrase the basic pre-
dicate is directly connected with any of its arguments, i.e. two
words of the primitive phrase cohere together when either the
first depends on the second, or the second on the first. When
illustrating the direct connectedness we can use the same diagramm
like when illustrating the dependence, but we do not pay attention
to arrow-heads. Thus, evidently in P "man" is directly connected

with "reads" but is not connected with "books" a.s.o.

But if there is given the phrase and if we have found all its primitive phrases, then it often happens that some word of the given phrase will occur (come up) in several primitive phrases. This evidently signifies that we are justified to identify the repeated words (naturally not always)occurring in different primitive phrases of the same sentence, and that is what we do by illustrating it on diagramm. In the considered case of the sentence "a man reads quickly a very good book" we get on to the diagram:

P

In this way, there are defined the semantical dependences among single words of the given sentence. The connectedness (no more direct) among single words of the sentence will be introduced in the way, common in the theory of graphes, namely, the words x is connected with the word y in the phrase P when there exists a finite sequence of the words of the sentence P, z_1, z_2, ..., z_K such that $z_1 = x$, $z_K = y$ and that z_i and z_{i+1} immediately cohere together for any $i = 1, 2, ..., k-1$.

Finally, if there is given the whole text consisting of single phrases, to which we have found their diagramms of the semantical dependence or connectedness, then often would happen that in the succeding phrases the same objects are spoken of, like in the preceeding-ones. Sometimes, this fact is distinctly expressed

by the referrin means (there are e.g. pronouns, definite articles and similar), but sometimes these are concealed and in this case it will be necessary to complete the text (or not to admit such a text at all). If there are everywhere the referring means expressed, they are possible to be used for further identification of the words of single diagramms for separate phrases (analogously as it was mentioned at the primitive phrases), and thereby to get the diagramms of the semantical dependence, eventually even the dependence for the whole text.

In the case of the whole-text-diagramm two cases are possible: either there is a connected graph and then we say that the connected text is concerned, or this graph is disconnected and then we say that the text is disconnected. But, any disconnected text splits, in a natural way, into its connected components and it is evident that it will be possible to translate these components independently on themselves (because they do not cohere together semantically).

Therefore we can concern only a connected context T. According to the section 2 $T = (S_1 . S_2 . \ldots S_k)$, where S_i are sentences and we remind that the condition (2) resp.(5) is not allways satisfied, because e.g. sometimes it is necessary to know, how the sentence S_1 was translated, when we want to translate correctly the sentence S_2. But now it is simple to see that there is exactely one word W^1 in S_1 and W^2 in S_2 such that W^1 and W^2 are directly connected. Therefore we may express a hypothesis that it is sufficient to store same informations concerning the single word W^1 instead of the whole translation of S_1.

In other words these informations concerning W^1 are the necessary context, when we want to translate conectly S_2. It is similar in other cases.

What concerns the translation of the particular sentences which are decomposed into the primitive phrases the main principales are described in [4] , because it is easy to indtroduced to each primitive phrase a corresponding rule as in a phrase structer grammar[1].

L I T E R A T U R E

[1] Chomsky N., On certain formal properties of grammars. Information and Control 2(1959),137-167

[2] Čulík K., Some problems in theory of langueges (Czech), Proceedings of 1.conference on cybernetics 1962 (in print)

[3] Čulík K., Application of abstract semantics and theory of graphs to polyglot dictioneries (Russian), Problems of Cybernetics (in print)

[4] Čulík K., Semantics and Translation of Crammars and ALGOL-like langueges,Kybernetika 1(1965),47-49

[5] Glushkoff V.M., Synthesis of digital automata (Russian), Moscow 1962