NEGATION AND FOCUS IN MAN - MACHINE COMMUNICATION

Petr Sgall, Eva Hajicova
Center of Numerical Mathematics
Charles University, Prague
Czechoslovakia

Abstract

A sketch of a programme for man-machine communication as being prepared by Prague linguists is presented, with a characterization of the scope of negation and of the nature of focus - two problems that appeared crucial with respect to the following isauesi (i) grammatical analysis of input texts, yielding unambiguous semantic representations (SR s) of subsequent sentences, (ii) algorithms translating the SR s into a logical language in the form of which the information gained from the natural language input is to be stored in the computer memory, (iii) form of the SR s, i.e. the language serving, among other purposes, as the output of (i) and input of (ii). It is shown especially that the logical impact of natural language negation can be accounted for only if the topic/focua dichotomy is duly analyzed.

1. A programme for man-machine communication, using a full-fledged analysis of natural language, is being prepared by linguists working at the Faculty of Mathematics and Physics, Charles University, Prague. The algorithms that are formulated are divided into the following groups:

(i) grammatical analysis of input texts (i.e. of Czech texts *on* electronics), yielding unambiguous semantic representations of subsequent sentences (in the form of linearized dependency graphs);

(ii) algorithms translating from the semantic representations (which still have a linguistic character) into a logical language, well adapted to the purposes of the theory *of* inference etc.; in this form the information gained from the natural language input is to be stored in the computer memory (and confronted with further such information, gained later from other texts);

(iii) algorithms looking for Information needed according to questions formulated by the users in Czech and analyzed according to the algorithms of (i);

(iv) algorithms of the synthesis of answers (in Czech).

We want to characterize in the present paper the solution of two problems that appeared crucial for (i) and (ii), as well as for the form of the semantic representations (SR 's) of sentences (the language serving as the output of (i) and input of (ii), among other purposes).

2. Linguistic considerations have led to the following form of SR's (where every A stands for a lexical unit, accompanied, possibly, by markers of "morphological meanings" (grammatemes), such as Plural, Preterite, etc., the subscript denoting the function of the given item in the sentence: Actor, Objective, Dative, Place, Direction, etc.):

(1) $V\ (A_1^b, A_2^b, \ldots, A_j^b, A_{j+1}, \ldots, A_n)$

Each of the participants, denoted here as A , can, of course, itself consist in a group headed by the verb, so that a structure of the shape (1) may be embedded, under certain conditions (specified by a generative grammar using dependency syntax) into another such structure.

The superscript b is attached to those elements that are contextually bound, i.e. included in the topic; the others constitute the comment or the focus of the sentence. The left-to-right order (with topic preceding the focus) may be regarded as a "deep" word order, corresponding to the primary word order of the surface level (combined with a neutral intonation). With a "hypersentence" of the Ross-Sadock type an 5R can be paraphrased as "I tell you about Topic that Focus", e.g. (<u>I tell you about</u>) <u>Father</u> (<u>that he</u>) <u>is coming</u>. (<u>I tell you</u>

about) reading newspapers (that it) can be boring. (I tall you about the fact) that Mabel came yesterday (that it) was surprising.

We consider the topic/focus dichotomy as relevant for the linguistic aspects of man-machine communication for the following reasons:

(a) The semantic content is organized in natural language in a different way than in formal logical languages, one of the main distinctions consisting in that what is rendered, e.g. in predicate logic, by the order of quantifiers, ia expressed in natural language (together with some other phenomena) by the articulation of the sentence into topic and focus and by the so-called scale of communicative dynamism (CD, realized, in the surface sentences, by certain features of word order and intonation, and denoted in our shape of SR' s by the left-to-right order of participants). As the exam - ales in the Appendix show, these phenomena are crucial for one of the moat basic discussions in contemporary linguistics (that between Chomskian and Lakovian wings in TG) and can be treated adequately if the topic/focus dichotomy is duly accounted for.

(b) An account of the dichotomy is not a mere queation of linguistic theory, but has a considerable impact on applications concerning the use of computers as already Garvin claimed [1] .

(c) winograd [2, p.116] wants to represent knowledge in an imperative form, so that not only the declarations as such, but also the way they are to be used should be specified. It must be noticed that in natural language (contrary to predicate logic) a declarative sentence is constructed not merely aa an assertion, but as an instruction to the hearer what familiar (known, already foregrounded in his memory) items he should pick up and modify somehow (combining them with new properties, or bringing them into new relations with other items, etc.). This introduction-like character of a sentence underlies its appropriateness to be used at a certain point of the discourse or text, or more broadly, in a certain situation; if we have Just spoken about George, we say George met Paul, while having spoken about Paul, we say rather Paul was met by George or It was George who met Paul. (In these simple examples, the two variants have identical truth conditions, but in the general case this does not hold, cf. the examplea in the Appendix.)

Our treatment of topic and focus, which follows, as for an empirical analysis, the tradition of Czech linguistics (relatively rich in this respect), is characterized in more detail in [3], where also a critical commentary on Chomsky s rather incomplete account of focus is given; it is shown there that operative and testable criteria such as the question test, the test of negation, properties of enclitical pronouns in some languages can be used to determine both the scale of communicative dynamism (which is an older but more precise formulation of what Chomsky calls "range of permissible focus") and the dichotomy of topic (Chomsky's and Jackendoff s "presupposition") and focus (the traditional comment or rheme).

3. Some important questions concerning SR s can be solved only if the position of negation is wall understood. It is well known today that meaning proper must be distinguished from presuppositions, or, more precisely, that three types of linguistic units must be distinguished) if the assertion A ia logically entailed by the given SR (in a certain situation and content), than:

(a) the negation of the given SR (Neg-SR) logically entails Neg-A; in this case, A belongs to the meaning proper of the sentences corresponding to that SR;

(b) Neg-SR logically entails A; here A is a presupposition:

(c) Neg-3R logically entails neither A, nor Neg-A; in this case, A is an allegation.

In such examples as (2) or (3) the entailed items (4) and (5), respectively, are allegations: thoy are not entailed by the negation of (2) or (3), and neither their negations are. On the other hand, if the relevant part of the sentence is contextually bound, the entailed item is presupposed; thus (6) presupposes (5), since not only (6) but also its negation entails (5). (The capital letters in the examples denote the bearer of the intonation centre.)

(c) he was upset by the arrival of my FRIENDS.

(3) John visited countries that suffered from poor CROPS.

(4) His friends CAKE.

(5) Some countries suffered from p#or CHOPS,

(6) Countries that suffered from poor crops were visited by JOHN.

An explanation offers itself in terms of the scope of negation: as our investigations of negation have shown, in the primary (unmarked) case, the scope of negation in specified on the left-hand side of the scale of CD by the position of the operator of negation and on the right-hand side by the end of the semantic representation of the sentence. The operator of negation stands either immediately before or immediately after the boundness juncture. {We use the term boundness juncture for the boundary between the contextually bound and the non-bound elementa of the SR,) primarily, the verb is non-bound (it predicates something about the topic), and the scope of negation extends from the boundnees juncture to the end of the semantic representation (He didn't COKE yesterday., cf. (7) below). Whet is negated here, is the relation between tne bound segment (the topic) and the focus. In the secondary case, where the verb is contextually bound, there are two possibilities: the negation of assertion with the scope of negation from the boundness juncture to the end of the semantic representation (It wasp' t YESTERDAY he came., cf. (8) below), or the negation of predication, where the scope of negation includes the verb alone (It was YESTERDAY he didn't come., cf. (9); the end of the scope of negation is the boundness juncture). Thus, the following three possibilities of the placement of the operator of negation and of the scope of negation with regard to the position of the boundness juncture are accounted for in the functional generative description:

$$(7)\quad V\ (A_1^b,\dots,A_j^b,\ Neg,\ A_{j+1},\dots,A_n)$$

$$(8)\quad V^b\ (A_1^b,\dots,A_j^b,\ Neg,\ A_{j+1},\dots,A_n)$$

$$(9)\quad V^b\ (A_1^b,\dots,A_j^b,\ Neg^b,\ A_{j+1},\dots,A_n)$$

where V is understood to stand - as for the scale of communicative dynamism - after all ccntextually bound elements (with an additional specification stating that Neg precedes V if the boundness juncture does not intervene). Notice that (7) ia the primary case, while (8) corresponds to the eo-called "constituent negation" in surface syntax, and in (9) only the verb ia negated (which corresponds to a "lexical negation" rather than to a negation of the SR as such. An allegation is entailed only by a sentence where the element determining it is not in the scope of negation (which does not necessarily coincide with the focus), but where the element determining an allegation is not connected with a corresponding presupposition* i.e. when the determining element is included in the scope of negation (as e.g. countries that ••• in the negation of (3)), the corresponding statement is no longer entailed. Meaning proper (assertion) consists in the relationship between what is talked about and what is said about it: in the sentence, the speaker asserts something (identified by the focus) about something (identified by the contextually bound segment); if the focus is negated, then the corresponding negative statement is entailed.

4. Taking into account also the delimiting features (such as Definite, Indefinite, Specifying), the shape of an SR as given in (1) should be complemented in the following way:

$$(10)\quad V_{m_0}\ ((P_1\ a_{m_1})^b_{k_1},\ (P_2\ a_{m_2})^b_{k_2},\dots$$
$$\dots,\ (P_j\ a_{m_j})^b_{k_j},\ (P_{j+1}\ a_{m_{j+1}})_{k_{j+1}},\dots$$
$$\dots,\ (P_n\ a_{m_n})_{k_n})$$

where the indices of the form m_i specify lexical units, the indices of the form k. specify the "cases" (participants, syntactic functions)[3] and P, stands for one of the operators SPEC, DEF, INDEF, EV, FEW, MANY, i.e. for delimiting features (understood here in the sense of Bierwisch[4] ; for a discussion cf [5] ; for the purpose of our present discussion, we class here also many and few as delimiting features). The formula (10) represents an SR of a sentence with a contextually non-bound verb; if the verb is contextually bound, then V is attached a superscript b..

We assume that it is possible to formulate a procedure translating the SR s to a formal language sharing the advantages of predicate calculus (with its rules of inference, etc.), but lacking some of its drawbacks (connected with the well known fact that the formulae of the calculus "do not say directly how they should be used" in procedures concorning inference). To this aim, we introduced a new operator St x (read ae: "for x such that ..."J. St is an operator with a free number of arguments; in its nature, it is close to the epsilon operator but it is generalized for a greater (arbitrary, but finite) number of name arguments having a set character. We write St(x)(F(x))(G(x))■ to be read "for such x for which F(x) holds, G(x) holds", which can be compared with Russell s G(x(F(x)))■ and *x* stands for a sequence of set name variables (where no ambiguity can arise, we do not dieting-uidh the name of an element from that of the set having only this one element).

The first step of this translating procedure, determined by Rule T, depends on the fact whether the verb occurring in the given SR is or is not contertually bound.

Rule T:

If the SR has the shape (10), it is rewritten as (11);

if the SR differs from (10) only in that V has the superscript b, the SR is rewritten as (12):

(11) $\mathrm{St}\,(x_{k_1}, x_{k_2}, \ldots, x_{k_j})\,(a_{m_1}(x_{k_1}^{P_1}) \,\&\ldots \&\, a_{m_j}(x_{k_j}^{P_j}))\,(\mathrm{St}\,(R, x_{k_{j+1}}, \ldots, x_{k_n})\,(V_{m_0}(R) \,\&\, a_{m_{j+1}}(x_{k_{j+1}}^{P_{j+1}}) \,\&\ldots\&\, a_{m_n}(x_{k_n}^{P_n}))\,(R(x_1, \ldots, x_n)))$

(12) $\mathrm{St}\,(x_{k_1}, x_{k_2}, \ldots, x_{k_j}, R)\,(a_{m_1}(x_{k_1}^{P_1}) \,\&\, a_{m_2}(x_{k_2}^{P_2}) \,\&\ldots\&\, a_{m_j}(x_{k_j}^{P_j}) \,\&\, V(R))\,(\mathrm{St}\,(x_{k_{j+1}}, \ldots, x_{k_n})\,(R(x_1, \ldots, x_n))\,(a_{m_{j+1}}(x_{k_{j+1}}^{P_{j+1}}) \,\&\ldots\&\, a_{m_n}(x_{k_n}^{P_n})))$

Note: It would be more correct if the superscript denoting the delimiting operator that binds the variable x_{k_i} ($i = 1, 2, \ldots, n$) were included also in the first parenthesis following the symbols St or R; but no misunderstanding can arise here.

Afterwards we eliminate, step by step, the occurrences of the symbol St (for a technical presentation of the whole procedure cf.[3, p.199-203], [6, § 10.2], and [7]) in such a way that it is replaced by one of the usual quantifiers (determined by the given delimiting feature) and, after other modifications stated in an algorithmic form, a formula of the second order predicate calculus is achieved (see the examples in the Appendix).

If an SR contains the symbol Neg, Rule T is applied first disregarding Neg; afterwards,

(a) if Neg has not been assigned the superscript b, then the symbol NON (for logical negation) is written to the left of the translation of the string $\Theta(\mathbf{v})$ in the formula representing the result of the application of T to the given SR;

(b) if Neg is present in the translated SR (before v^b, then NON V instead of V is written in the resulting formula.

Appendix: [4]

Example 1:

(a) A boy made a CANOE out of every log.

(b) $\mathrm{Make}^b\ ((\mathrm{SPEC\ Boy})_1^b,\ (\mathrm{EV\ Log})_3^b\ ,\ (\mathrm{SPEC\ Canoe})_2)$

(c) $\mathrm{St}(x_1, x_3 R)\ (\mathrm{Boy}(x_1^{SPEC}) \,\&\, \mathrm{Log}(x_3^{EV}) \,\&\, \mathrm{Make}(R))\ (\mathrm{St}(x_2)\ (R(x_1, x_2, x_3))\ (\mathrm{Canoe}(x_2^{SPEC})))$

(d_1) $\exists\, x_1(\mathrm{Boy}(x_1) \,\&\, \mathrm{St}(x_3, R)\ (\mathrm{Log}(x_3^{EV}) \,\&\, \mathrm{Make}\ (R))\ (\mathrm{St}(x_2)\ (R(x_1, x_2, x_3))\ (\mathrm{Canoe}(x_2^{SPEC}))))$

(d_2) $\exists\, x_1 \forall x_3(\mathrm{Boy}(x_1) \,\&\, \mathrm{Log}(x_3) \,\&\, \mathrm{St}(R)\ (\mathrm{Make}(R))\ (\mathrm{St}(x_2)\ (R(x_1, x_2, x_3))\ (\mathrm{Canoe}\ (x_2^{SPEC}))))$

(d_3) $\exists\, x_1\ \forall\, x_3\ \exists\, R(\mathrm{Boy}(x_x) \,\&\, \mathrm{Log}(x_3) \,\&\, \mathrm{Make}(R) \rightarrow \rightarrow \exists\, x_2\ (R(x_1, x_2, x_3) \rightarrow \mathrm{Canoe}(x_2)))$

Example 2:

(a) The boy told few girls about many PROBLEMS.

(b) $\mathrm{Tell}\ ((\mathrm{DEF\ Boy})_1^b,\ (\mathrm{FEW\ Girl})_2,\ (\mathrm{MANY\ Problem})_3)$

(c) $\mathrm{St}(x_1)\ (\mathrm{Boy}(x_1^{DEF}))\ (\mathrm{St}(R, x_2, x_3)\ (\mathrm{Tell}(R) \,\&\, \mathrm{Girl}(x_2^{FEW}) \,\&\, \mathrm{Problem}(x_3^{MANY}))\ (R(x_1, x_2, x_3)))$

(d_1) $\exists! \, x_1 (Boy(x_1)\ (St(R,x_2,x_3)\ (Tell(R)\ \&$

$\& \ Girl(x_2^{FEW})\ \&\ Problem(x_3^{MANY}))\ (R(x_1,x_2,x_3)))$

(d_2) $\exists! \, x_1 (Boy(x_1) \rightarrow \exists R (Tell(R)\ \&\ St(x_2,x_3)$

$(Girl(x_2^{FEW})\ \&\ Problem(x_3^{MANY}))\ (R(x_1,x_2,x_3))))$

(d_3) $\exists! \, x_1 (Boy(x_1) \rightarrow \exists R\ \exists y\ \forall x_2 (Tell(R)\ \&$

$\&\ y \subseteq Girl\ \&\ (x_2 \in y \rightarrow \exists R'(R' \in FEW \rightarrow$

$\rightarrow R'(y)))\ \&\ St(x_3)\ (Problem(x_3^{MANY})$

$(R(x_1,x_2,x_3)))))$

(d_4) $\exists! \, x_1 (Boy(x_1) \rightarrow \exists R\ \exists y\ \forall x_2\ \exists y'\ \forall x_3$

$(Tell(R)\ \&\ y \subseteq Girl\ \&\ (x_2 \in y \rightarrow \exists R'$

$(R' \in FEW \rightarrow R'(y)))\ \&\ (y' \subseteq Problem\ \&$

$\&\ (x_3 \in y' \rightarrow \exists R''(R'' \in MANY \rightarrow R''(y))))) \rightarrow$

$\rightarrow R(x_1,x_2,x_3))$

EXAMPLE 3:

(a) The girl saw no BOY.

(b^1) $See((DEF\ Girl)_1^b,\ Neg,\ (INDEF\ Boy)_2)$

(c^1) $St(x_1)\ (Girl(x_1^{DEF}))\ NON\ (St(R,x_2)\ (see\ (R)\ \&$

$\&\ Boy(x_2^{INDEF})\ (R(x_1,x_2,x_3))))$

(d^1) $\exists! \, x_1 (Girl(x_1) \rightarrow NON\ \exists R\ \exists x_2 (See(R)\ \&$

$\&\ Boy(x_2) \rightarrow R(x_1,x_2,x_3)))$

(b^2) $See^b((DEF\ Girl)_1^b,\ Neg,\ (INDEF\ Boy)_2)$

(c^2) $(St(x_1,R)\ (Girl(x_1^{DEF})\ \&\ See(R))\ NON\ (St(x_2)$

$(R(x_1,x_2,x_3))\ (Boy(x_2^{INDEF})))$

(d^2) $\exists! \, x_1\ \exists R (Girl(x_1)\ \&\ See(R) \rightarrow NON\ \exists x_2$

$(R(x_1,x_2,x_3) \rightarrow Boy(x_2)))$

(b^3) $See^3((DEF\ Girl)_1^b,\ Neg^b,\ (SPEC\ Boy)_2)$

(c^3) $St(x_1,R)\ (Girl(x_1^{DEF})\ \&\ NON\ See(R))\ (St(x_2)$

$(R(x_1,x_2,x_3)))\ (Boy(x_2^{SPEC})))$

(d^3) $\exists! \, x_1\ \exists R (Girl(x_1)\ \&\ NON\ See(R) \rightarrow$

$\rightarrow \exists x_2 (R(x_1,x_2,x_3) \rightarrow Boy(x_2)))$

Note: In Example 3 we illustrate the three positions of Neg by giving three different SR's, namely (b^1), (b^2), and (b^3). Since the delimiting feature INDEF cannot occur in (b), i.e. out of the scope of negation after auch a verb as "see", we work instead with SPEC.

Notes:

See e.g. [8,pp .174f.] and [9] for the notion of preeupposition; [10] and [11] for the trichotomy of semantic units as discusaed below.

2

For a detailed discussion of the motivations that led to the specification of the scope of negation as summarized here, cf. [12] • Various operational tests can be used to check the scale of CD, the contextual boundness and the position of negation in an SR; the well-known question test, used in our investigations, has yielded results similar to those reached by [I3, esp. Chapter 2] ; cf. also [14] .

The double indices k_i stand for permutations of the numbers 1,..,n, not necessarily ordered according to the systemic ordering, cf. what was said about topicalization in [3] s but this ordering necessarily appears inside the translated formula, cf. the substring $R(x_1,\dots,x)$ in (1) and (2). It is possible to introduce also a more subtle quantification of verbs (for repeated actions, non-immediate or gnomic events, etc.), cf. [15] and [16]

4

Our examples are chosen in such a way that they might show how the proposed type of description permits to avoid the difficulties that have forced Chomsky to "extend" his standard theory, and Lakoff to introduce the (too strong) global constraints. - We do not specify the morphological gramnatemes (of tense, number, etc.) here, since they are not relevant for the present discussion. Under (a) we give here the surface sentence {capital letters denoting the bearer of the intonational centre), (b) being one of its SR s, (c) the result of the application of T (and its extension for SR s containing the symbol Neg), (d) the reault(s) of (individual steps of) further algorithm yielding finally a formula of the second order predicate calculus.

Referencos

[1] P.Garvin, Fulcrum Techniques to Semantic Analysis, Final Report, Thompson Ramo Wocldridge Inc., Canoge Park, Calif-, 1963

[2].Winograd, Understanding Natural Language, Edinburgh 1972

[3] P.Sgall, E.Hajicova, and E.Benesova, Topic, Focus, and Generative Semantics, Kronberg/Ts., 1973

M.Bierwiech, On Classifying Semantic Features, in: Semantics, ede. D.D.Steinberg and L.A.Jakobovite, Cambridge 1971, pp. 410-435

[4] P.Krizek, Towards a Formal Account of the Semantics of Noun Phrases, Prague Bulletin of Mathematical Linguistics 20, 1973, pp. 43-58

E.Benesova, E.Hajicova, and P.Sgall, Hemarks on Topic/Comment Articulation, Part I, Prague Bulletin of Mathematical Linguistics 19, 1973, pp. 29-58; Part II, PBML 20, 1973, pp. 3-42, see eap. $ 10.2

[7] E.Hajicova, P.Kriiek, and P.Sgall, A Generative Approach to Semantica, in press in Linguistica Silesiana, Katowice

P.F.Strawaon, Introduction to Logical Theory, London 1952, pp. 174f.

C.J.Fillmore, Types of Lexical Information, ins Studies in Syntax and Semantics, ed. by F.Kiefer, Dordrecht 1969, pp. 109-137

[10] E.Hajicova, Some Hemarks on Presuppositions (presented at 1971 International Conference On Computational Linguistics, Debrecen), Prague Bulletin of Mathematical Linguistics 17, 1972, pp. 11-23

E.Hajidova, Meaning, Presupposition, and Allegation, Philologies Praeensia 17, 1974, pp. Iti-25

E.hajidova, Negation and Topic vs. Comment, Philologica Pragensia 16, 1972, pp. 81-93

R.Posner, Theorie des Kommentierens, Frankfurt/m. 1972, (esp. Chapter 2)

[14] J.M.Zemb, Las structures logiques da la proposition, allemande, Paris 1969

Y.Bar-Hillel, Indexical Expressions, Mind 63, 1954, pp. 359-371

[16] D.Davidson, The Logical Form of Action Sentences, in: The Logic of Decision and Action ed. N.Rescher , Pittsburg 1968, pp. 81-120.