Topologies and Sheaves in Discourse Analysis

OLEG PROSOROV

St. Petersburg Department of V. A. Steklov Institute of Mathematics

This talk presents a sheaf-theoretic formal semantics as a novel paradigm in discourse analysis. We consider mainly a written type of linguistic communication in some unspecified natural language, say English, French, Russian, and so its basic units are texts. All the texts we consider are supposed to be written with good grace and intended for human understanding; we call them *admissible*. A sentence is considered as a sequence of its words and a text as a sequence of its sentences. Any part of a considered whole is simply a subsequence of a given sequence.

1. Basic Concepts. – We distinguish the semantic notions *sense*, *meaning* and *reference*. The term *fragmentary meaning* of some fragment of a given text is accepted as the content grasped in some particular situation of reading guided by the reader's presuppositions and preferences in the interpretative process, which we denominate by the term *sense* (or *mode of reading*). This *sense* is a kind of semantic orientation in the interpretative process that relates to the totality of text or its fragment, sentence or its syntagma, and involves the reader's subjective premises that what is to be understood constitutes a meaningful whole; it is in some degree a secular remake of the term *sense* in the medieval exegesis (St. Thomas).

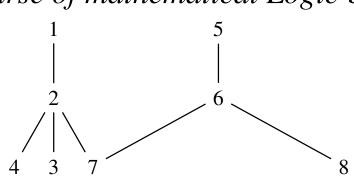
2. Phonocentric Topology. – In the process of reading, the understanding is not postponed until the final sentence of a given text. So the text should have the meaningful parts and the meanings of these parts determine the meaning of the whole as it is postulated by the principle of *hermeneutic circle*. It seems to be quite in agreement with our linguistic intuition that:

(i) *an arbitrary union of meaningful parts of an admissible text is meaningful;*
(ii) *a non-empty intersection of two meaningful parts of an admissible text is meaningful.*

For an admissible text is supposed to be meaningful as a whole by definition, it remains only to define the meaning of its empty part (e.g. as a one-element set) in order to provide it with some *topology* in a strict mathematical sense, where the open sets are all the meaningful parts (called further *fragments*). We call *phonocentric* the topology so defined. For every pair of different sentences x, y of a text X, there exists an open $U \subset X$ containing precisely one of them; whence a phonocentric topology should satisfy the *separation axiom* T_0 of Kolmogoroff. Another concept of meaning or criterion of meaningfulness let us to define yet another type of topology on X. An admissible text X gives rise to a finite space, hence an arbitrary intersection of its open sets is open and so it is an *Alexandroff space*. For a sentence $x \in X$, we define U_x to be the intersection of all the meaningful parts that contain x, i.e. the smallest open neighborhood of x. We define the *specialization* relation $\preceq$ on X by setting $x \preceq y$ if and only if $x \in U_y$ or, equivalently, $U_x \subset U_y$. Note that for all $x, y \in X$, $x \preceq y$ implies $x \leq y$, where $\leq$ defines the natural linear order of reading.

Proposition. *The set of all open sets of the kind U_x is a basis of a phonocentric topology on X. Moreover, it is the unique minimal basis of a phonocentric topology. The phonocentric topology on an admissible text defines a partial order structure $\preceq$ on it by means of specialization; the initial phonocentric topology can be recovered from this partial order $\preceq$ in a unique way.*

There exists a simple intuitive tool called Hasse diagram for the graphical representation of a finite partially ordered set (*poset*). For a poset $(X, \preceq)$, the *cover relation* $\prec$ is defined by: '$x \prec y$ if and only if $x \preceq y$ and there exists no element $z \in X$ such that $x \preceq z \preceq y$'. In this case, we say that *y covers x*. For a given poset $(X, \preceq)$, its Hasse diagram is defined as the graph whose vertices are the elements of X and whose edges are those pairs $\{x, y\}$ for which $x \prec y$. The usage of some kind of Hasse diagram under the name of *Leitfaden* is widely spread in the mathematical books. So, in *A course of mathematical Logic* by Yu. Manin, it appears like this:



In this usage, the poset is constituted of all chapters of the book. The picture may be "split" in order to draw the Hasse diagram whose vertices are all the paragraphs, and so on. These considerations may be repeated with slight modifications in order to define a *phonocentric topology* at the semantic level of sentence and even word. Thus for a given admissible text, we can find, in a constructive manner, its phonocentric topology at each semantic level. Now we may interpret linguistic notions in terms of topology and order and undertake its geometrical studies, which may be thought of as a kind of *formal textual syntax*.

3. Sheaves of Fragmentary Meanings. – Let X be an admissible text, and let $\mathscr{F}$ be an adopted *sense* or *mode of reading*. For a given fragment $U \subset X$, we collect all the fragmentary meanings of U in the set $\mathscr{F}(U)$. Thus we are given a map $U \mapsto \mathscr{F}(U)$ defined on the set $\mathscr{O}(X)$ of all opens $U \subset X$. Formulated not only for the whole text X but more generally for any meaningful part $V \subset X$, the precept of the hermeneutic circle 'to understand any part of text in accordance with the understanding of the whole text' defines a family of maps $\mathrm{res}_{V,U} \colon \mathscr{F}(V) \to \mathscr{F}(U)$, where $U \subset V$, such that $\mathrm{res}_{V,V} = \mathrm{id}_{\mathscr{F}(V)}$ and $\mathrm{res}_{V,U} \circ \mathrm{res}_{W,V} = \mathrm{res}_{W,U}$ for all opens $U \subset V \subset W$. From a mathematical point of view, the data $(\mathscr{F}(V), \mathrm{res}_{V,U})_{V,U \in \mathscr{O}(X)}$ is a *presheaf* of sets on X.

The reading process of a given fragment U is modeled as its (open) covering by some family of subfragments $(U_j)_{j \in J}$, where each U_j is supposed to be read in a distinct physical act.

According to Quine, there is no entity without identity. The definition of equality that seems to be quite adequate to our linguistic intuition is posed by the following:

Claim S (Separability). *Let X be an admissible text, and let U be a fragment of X. Suppose that $s,t \in \mathscr{F}(U)$ are two fragmentary meanings of U and there is an open covering $U = \bigcup_{j \in J} U_j$ such that $res_{U,U_j}(s) = res_{U,U_j}(t)$ for all fragments U_j. Then $s = t$.*

Thus an adopted sense (mode of reading) of a text X determines some *separated presheaf* $\mathscr{F}$ of its fragmentary meanings. Following the precept of the hermeneutic circle 'to understand the whole text by means of understandings of its parts' this presheaf $\mathscr{F}$ should satisfy the following:

Claim C (Compositionality). *Let X be an admissible text, and let U be a fragment of X. Suppose that $U = \bigcup_{j \in J} U_j$ is an open covering of U; suppose we are given a family $(s_j)_{j \in J}$ of fragmentary meanings, $s_j \in \mathscr{F}(U_j)$ for all fragments U_j, such that $res_{U_i, U_i \cap U_j}(s_i) = res_{U_j, U_i \cap U_j}(s_j)$. Then there exists some meaning s of the whole fragment U such that $res_{U,U_j}(s) = s_j$ for all U_j.*

A separated presheaf satisfying the claim (C) is called a *sheaf*. It imposes the following:

Definition (Frege's Generalized Compositionality Principle). *A separated presheaf of fragmentary meanings naturally attached to any sense (mode of reading) of an admissible text is really a sheaf; its sections over any fragment of the text are the fragmentary meanings; its global sections are the meanings of the whole text.*

Recall that the elements of $\mathscr{F}(U)$ are usually called *sections of $\mathscr{F}$ over U*. We note that the claim (S) guarantees the meaning s, whose existence is claimed by (C), to be unique as such.

A *morphism* of sheaves $\phi \colon \mathscr{F} \mapsto \mathscr{F}'$ on the same text X is a family of maps $(\phi(V))_{V \in \mathscr{O}(X)}$, where each $\phi(V) \colon \mathscr{F}(V) \to \mathscr{F}'(V)$ represents a transfer from the understanding of V in the sense $\mathscr{F}$ to its understanding in the sense $\mathscr{F}'$ which is compatible with the restriction maps, i.e. $\phi(U) \circ \mathrm{res}_{V,U} = \mathrm{res}'_{V,U} \circ \phi(V)$ for all $U \subset V$.

Thus, given an admissible text X, the data of all sheaves $\mathscr{F}$ of fragmentary meanings together with all its morphisms constitutes a category $\mathbf{Schl}(X)$, called *category of Schleiermacher*.

4. Sheaf-theoretic formal semantics. – This approach describes a natural language in the category of *textual spaces* **Logos**. The objects of this category are couples $(X, \mathscr{F})$, where X is a topological space naturally attached to an admissible text and $\mathscr{F}$ is a sheaf of fragmentary meanings on X; any such a couple is called a *textual space*. The morphisms are couples $(f, \theta) \colon (X, \mathscr{F}) \to (Y, \mathscr{G})$ made up of a continuous map $f \colon X \to Y$ and a f-morphism of sheaves θ which respects the concerned sheaves, i.e. $\theta \colon \mathscr{G} \to f_* \mathscr{F}$, where f_* is a *direct image* functor.

The different categories and functors related to discourse and text interpretation processes are the principal objects of study in the *sheaf-theoretic formal semantics* as we understand it.

References

Prosorov, O. (2005a). Compositionality and contextuality as adjoint principles. In M. Werning, E. Machery, & G. Schurz (Eds.), *The compositionality of meaning and content, (Vol. II : Applications to Linguistics, Psychology and Neuroscience)* (pp. 149–174). Frankfurt: Ontos-Verlag.

Prosorov, O. (2005b). Formal hermeneutics based on Frege duality. In E. Maier, C. Bary, & J. Huitink (Eds.), *Proceedings of SuB9* (pp. 286–298). Nijmegen: NCS.

Prosorov, O. (2006). Sheaf-theoretic formal semantics. *TRAMES Journal of the Humanities and Social Sciences*, *10*(1), 57–80. (Published by the Estonian Academy of Sciences and Tartu University)

Tennison, B. R. (1975). *Sheaf theory*. Cambridge: Cambridge University Press.