Topology as textology. How semantic diagrams can contribute to a modern hermeneutics

Timon Georg Boehm[0000−0001−5318−6871]

International Nietzsche Research Group at the University of Stuttgart
philosophie@timonboehm.ch

Abstract. In recent decades the digital revolution has strongly influenced paradigms and methods in the humanities. Particularly for literary studies, philology, or philosophy this raises the question of what a contemporary hermeneutics should look like that keeps pace with this development instead of simply surrendering to it. The main goal of this paper is therefore to provide new concepts to extend traditional hermeneutics in a way that takes into account modern (often quantitative) approaches and still complies with its former achievements. To accomplish this, the potential of mathematical topology is explored which promises several advantages at this point: i) its concepts are both formally usable and intuitively understandable, ii) contexts are always already taken into account, iii) semantics is not only based on material characters but on extended regions in topological spaces. Thus semantical structures can be described by different topologies and those topologies be represented as diagrams.

Keywords: Topology, Graphtheory, Hermeneutics

1 State-of-the-art

It has become popular to collect and analyse linguistic data with a wealth of programmes. A persistent problem of such quantitative text analyses, however, is the inclusion of contexts and their interconnectedness. This is due to the fact that usually information of a text is extracted in the form of word frequencies (and stored e. g. as a document feature matrix). All further analysis is based then on this data such that ultimately meaning must be reconstructed from numbers. A concrete problem when neglecting contexts is, for instance, the ambiguity of words. "Polysemy is a tough problem for distributional approaches, as a representation that is learned from all of a word's contexts will conflate the different senses of the word." ([1], p. 635). Convincing solutions are lacking so far.

2 Objectives

The general aim of the novel "Topology as textology" approach is to describe meaning-constitutive structures in texts by concepts of mathematical topology and to establish thereby elements of a modern hermeneutics. In particular, this paper addresses the following questions:

- How the mathematical topological vocabulary (such as space, neighbourhood, intersection, union, connection, path, region, boundary etc.) has the potential to capture meaning and sense in texts in a way that is compatible with modern (mostly quantitative) text studies and, simultaneously, can be docked onto hermeneutical concepts.
- How contexts are generically included by using the mathematical topological vocabulary, due to the fact that in mathematical topology all entities are always already determined by their belonging to a whole and do not have to be artificially assembled.
- How topology can accommodate for semantic spaces in the full sense and therefore satisfy hermeneutical exigences. This idea differs from existing approaches under the title of 'semantic topology' in that it avoids any reductionism, be it on the quantitative, be it on the qualitative side.

3   Methodology

The study is conducted on two sample texts, namely *Human, all too Human 1* and *Twilight of Idols* by Friedrich Nietzsche. The starting point is the observation that "the central link between traditional hermeneutic textual approaches and algorithmic text analysis is operationalisation" (Pichler and Reiter, in prep.). Such an operationalisation shall be done first through graph-theoretical representations of linguistic terms, putting single terms or whole text-units in relation to each other. Such graphs are abstractions from the original text where all information is captured, amongst others, by the neighbourhood of a node (the set of vertices connected to it by an edge) or by the connectivity of the graph. Since this can be done by fairly standard programmes for corpus analysis (such as the free software 'R' which is maintained by a professional community, www.r-project.org), this will anchor the new method to existing methods.

Graphs, however, rely on material text-elements and do not yet span semantic spaces in the full sense. One can reach, and this is the key hypothesis, a truly semantic level by embedding graphs in topological spaces which is also a mathematically well-defined operation. Each material term or text-unit is thus assigned a neighbourhood or region in a semantic space. The characteristic connection of such neighbourhoods or regions is then a topology. Thus texts are semantic topological spaces, where 'space' is meant mathematically and 'semantic' hermeneutically.

4   Preliminary results

To give a simple example of a preliminary study for this new methodology: Consider how a chapter of *Twilight of Idols* is composed of different preliminary stages in the notebooks and draw a graph which represents this composition. Each node corresponds to a preliminary stage, each edge corresponds to the boundary from one preliminary stage to the other in the finished text. Denoting the workbooks with the common Mette-signature, page number and line

number, we have e. g.: WII6,104,1-18 $\rightarrow$ WII6,105,6-10 $\rightarrow$ WII6,104,20-24 $\rightarrow$ WII6,105,16-42 $\rightarrow$ WII7,38,2-34 $\rightarrow$ .... If these nodes and edges are embedded in a topological space, the boundaries in the text become the boundaries of semantic regions, in this case 'physiology', 'psychology', 'moral philosophy', 'metaphysics'. Nietzsche's argument can then be understood by considering their characteristic connection, i. e. topology. Possible candidates here are the so-called discrete topology $\mathcal{T} = \{\emptyset, \{p\}, \{q\}, \{r\}, \{p,q\}, \{p,r\}, \{q,r\}, \{p,q,r\}, \ldots\}$ or the nested topology $\mathcal{T} = \{\emptyset, \{p\}, \{p,q\}, \{p,q,r\}, \ldots\}$, reflecting the fact that moral philosophy is to be investigated on the ground of psychology and the latter on the ground of physiology. The same procedure can be applied to aphorism books or, in fact, any text with suitable units. Like this, also ambiguities can be resolved by attributing them to different topologies. The idea to apply topologies to texts has been put forward already by Max Bense, but he limited it to the material level of single words whereas, and this is a novelty of the TopText project, the full strength of topology only comes to bear at the semantic level. Note that the existence itself of bounded semantic regions in a text cannot just be read off but is a transcendental assumption for understanding at all.
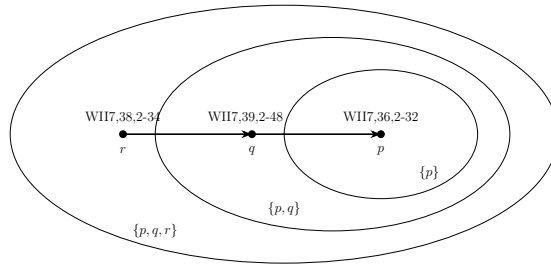
Fig. 1. Section from the embedding of a graph in a topological space. The three domains or meaning components are open sets as required by axiomatics.

If these preliminary results are extended to the full text we can reconstruct a genetic line of argument and represent it as a diagram. Of course, the example diagrams here are very simple in order to explain the method in principle. In reality, we have to consider a great number of nodes, edges, and semantic regions which can be handled only through a meaningful representation. Diagrams are essential to the method because the semantical structure of a text is mapped into a spatial structure and therefore its semantics is understood precisely through diagrams.

References

1. Erk, K.: Vector Space Models of Word Meaning and Phrase Meaning: A Survey. Language and Linguistics Compass **6**(10), 635–653 (2012)