THE SOCIAL NETWORK OF DANTE'S *INFERNO*

Amedeo Cappelli, Michele Coscia, Fosca Giannotti, Dino Pedreschi, Salvo Rinzivillo, KDDLab ISTI-CNR Pisa, PI, Italy
E-mail: <name.surname@isti.cnr.it>

Submitted: April 29, 2010

Abstract

Our work aims to approach the phenomenon of culture through the development of new methods and more powerful tools to capture the content of digitally stored literary material. As our test bed we choose Dante's characters of *al di là*, a domain consisting in a set of data and relations complex enough to sharpen our tools. The methods of investigation we intent to realize should help scholars of a literary text to concretize part of his interpretative intentions. For this purpose, we will adopt advanced methodologies suitable to obtain a rich representation, expressing the multi-relational network inherent in the text.

Keywords: Complex Networks, Dante, Multirelation, Divina Commedia, Ontology, Inferno, Visual Analytics

A literary text is a complex organism in which many logical and knowledge entities interact both in an implicit and in an explicit way. In particular, new methods and more powerful tools are needed to capture the content of digitally stored literary material. We do believe that complex network analysis and visualization can provide some of these methods. For demonstration, we chose as case study Dante's characters of *al di là*, a domain consisting of a set of data and relations so complex to be adequate to match the above requirements. It is well known among Dante scholars that the interpretation of the *Commedia* requires the evaluation of many types of knowledge regarding both the structure of the text as well as the general context the text has been generated. The context includes the whole set of historical, political, social, philosophical and cultural experiences of Dante, which have been explicitly or implicitly involved in the production of his *Commedia*.

We adopt both analytical and visual methodologies suitable to obtain a rich representation of the textual content, extracting and depicting what we call a multidimensional social network. A social network is a structure in which two social entities are connected by links according to one particular criterion [1]. In a multidimensional, or multi-relational, network the links are typed according to particular disjoint criteria.

System Specifications

A system able to analyze and interpret a text in a relational way, must adapt itself to scholarly behavior, by expressing informative intentions regarding many aspects of a text. Consequently, such a system must exhibit the necessary expressive power able to analyze and understand complex multi-relational problems a scholar may formulate. This may require the retrieving and the representation of a vast set of interacting concepts in the framework. And it may lead to a global representation of micro and macro phenomena which actually represent the inner structures of a text – presumably chosen by an author consciously, or on the basis of external social constraints.

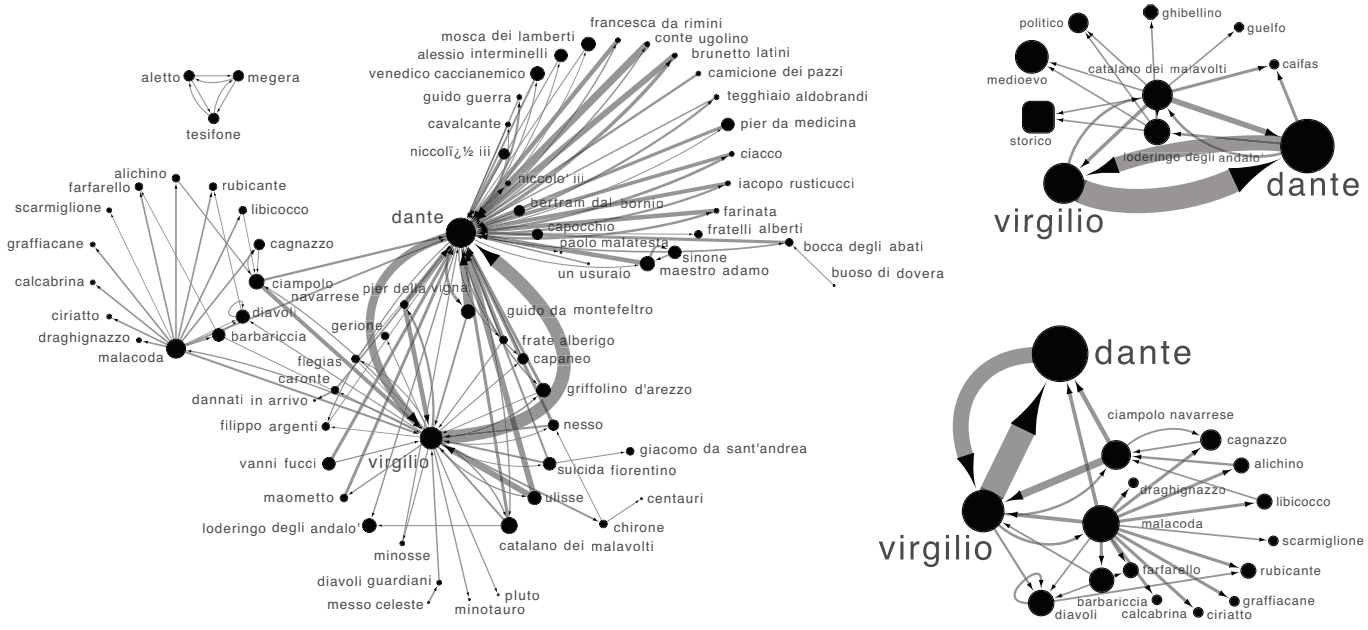
It furthermore requires that the network building procedure is able to produce structured results consisting of interacting entities in multi-relational network form, computed with respect to all available data, by far surpassing what can be held in the memory of a human scholar. In addition it is clear that the latent knowledge emerging from the analysis of the complex informative structure of the text is incomplete. Many preconditions to the interpretation of the *Commedia* are external to Dante's text. It is necessary to take the knowledge on Dante and his time into consideration for a full interpretation. Our method tries to consider both the structure of the text and part of the external knowledge in a two level structure. Doing so, new complex paths of interacting concepts emerge naturally, providing different points of view towards the literary data.

To this aim, we represent the text as a multi-dimensional or multi-relational network, including at different levels, the literary context together with a suitable representation of the inherent context – historical, social, political, etc. Although the present paper focuses on social examples, the same analysis can be performed including any or all other dimensions.

Building the network

We build the respective network using both the dialog relation (i.e. who talks to whom) and several further relations based on knowledge about external information used by Dante. We end up with a network in which nodes are

Fig. 1. Visualization of the *Inferno* social network, along with two conceptual ego networks, extracted from the main network and ontology data.



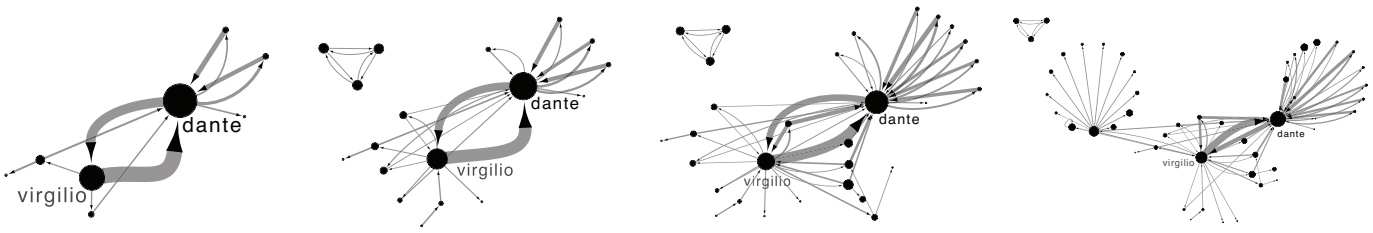


Fig. 2. Temporal evolution of the *Inferno* social network within the narrative timeline.

acters, connected if they speak to each other, and concepts connected to relevant characters. These relations were extracted from a manually annotated version of the *Inferno*, according to a process explained in [2] and [3], where relevant information about the speaking characters is collected and linked to the dialog by a domain expert, using the social, political, historical, mythological and linguistic contexts.

In the main social network the entities are the characters of the *Inferno*. A weight is assigned to link according to the number of verses in which one character talks to another. The whole social network of the dialogues is represented in the left part of Fig 1.

There are two additional information levels attached to this social network. First, we have "temporally annotated" the connections according to the part of the *Inferno* in which the dialogue appears. Subsequently it is possible to visually grasp the evolution of the flow of information present in the *Inferno*. Fig. 2 gives an impression of this dynamic evolution, showing four snapshots corresponding to four of the original divisions in the *Canti* – VI, X, XVI, and XXII.

Second, we have linked each character to concepts in an *ontology*. For example, consider the character Catalano dei Malavolti, an historical character that is a medieval guelph politician present in the *Cerchio* chapter VIII. Therefore, the character node "Catalano dei Malavolti" is connected to the concept node "Historical" in the ontology using the link "character type", to "Medieval" in using the "historical period" link, and to "Guelph" using the link type "political faction", etc.

With this additional information we enable users, who are interested in studying the interplay among several different concepts and characters, to navigate and explore such relations. Initially, we visualize the main social network without the ontology information. Then the user can select one particular character and focus on all the concepts related to him, restricting the visualization to immediately relevant concepts and further characters sharing the very same concepts. Since these networks are centered on a particular entity, we call them "ego networks" [4]. The process is depicted symbolically in Fig 1, where in the right part we show two possible results of two different selections: the ego network of Malacoda and Catalano dei Malavolti.

Results

As stated in the system specification section, the main focus of our work is to create a novel point of view on the interacting concepts of a literary text. We have described a system able to create an immersive knowledge environment. Inside this environment, the scholar is surrounded by a complex, multi-dimensional relation space, defined by all the concepts the author included, implicitly and explicitly in his work, as far as they are made explicit by the curators who annotated the text. The system provides a browsing feature able to focus the attention of the scholar on a particular topic. With this system, one could address some questions like: Why does Dante use a dialectal lexicon for a particular character?, Are Guelphs and Ghibellines randomly distributed in the text or are there some particular patterns in their appearances?, etc.

Furthermore, we are confident that some non-trivial analytical results could also be achieved. Consider Fig. 2, where according to the visible the evolution one can say that the network exhibits a "winner takes all" effect – very similar to the well known *preferential attachment* model of the Web [5]. The two main hubs, Dante and Virgilio, and their dense communication links, reflect the central role that Dante has reserved for Virgilio, representing the *magister* that drives Dante through the *al di là*.

Fig. 3. The workflow of our work: from the plain text to the complex network.

References and Notes

* This paper was presented as a contributed talk at Arts | Humanities | Complex Networks – a Leonardo satellite symposium at NetSci2010. See <http://artshumanities.netsci2010.net>

1. M. E. J. Newman, The structure and function of complex networks. (*SIAM Review* 45, 167-256, 2003).

2. A. Cappelli, M. N. Catarsi, P. Michelassi, L. Moretti, M. Baglioni, F. Turini, M. Tavoni, Knowledge mining and discovery for searching in literary texts. (*Proceedings of the Third International Conference on Language Resources and Evaluation*, Las Palmas, 2002).

3. Miriam Baglioni, Mirco Nanni, Emiliano Giovannetti, Mining Literary Texts by Using Domain Ontologies. (*Proceedings of the Knowledge Discovery and Ontologies workshop in ECML/PKDD*, Pisa, 2004).

4. Linton C. Freeman, Centered graphs and the construction of ego networks. (*Mathematical Social Sciences* Volume 3, 1982).

5. Albert-Laszlo Barabasi, Reka Albert, Emergence of scaling in random networks, (*Science* 286, 509, 1999).