POLITECNICO DI TORINO
Repository ISTITUZIONALE

Using Time Series and Graphs in the Analysis of Dante's Divine Comedy

Original
Using Time Series and Graphs in the Analysis of Dante's Divine Comedy / Sparavigna, A.C.. - In: INTERNATIONAL JOURNAL OF SCIENCES. - ISSN 2305-3925. - STAMPA. - 3:12(2014), pp. 33-40. [10.18483/ijSci.605]

Availability:
This version is available at: 11583/2583558 since: 2015-09-03T06:42:49Z

Publisher:
England: Alkhaer Publications

Published
DOI:10.18483/ijSci.605

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

12 September 2026

International Journal of Sciences

Research Article Volume 3 - December 2014 (12)

Using Time Series and Graphs in the Analysis of Dante's Divine Comedy

Amelia Carolina Sparavigna[1]

[1]Department of Applied Science and Technology, Politecnico di Torino, Italy

Abstract: The digitized form of a literary work can be used to have a numerical analysis of it, for instance, by giving the frequencies of words occurring in the text and a consequent appraisal of characters, places or leitmotivs to which these words are linked. The related numerical data can also be organized and analysed as time series, where the time is played by the progression of lines of the digitized text. Here we propose examples of time series concerning some words in Dante Alighieri's poem, the Divine Comedy. Besides plotting the time series, we can also use graphs for illuminating some leitmotivs of this poem.

Keywords: Literary Experiments, Word Occurrences, Graphs, Divina Commedia, Divine Comedy, Dante Alighieri

1. Introduction

In some recent papers [1-3], we have discussed the use of network theories in the analysis of characters in plays and novels. The use of network theories for literary works is quite interesting [4-7], but a problem exists: in this approach the evolution of their plots is lost, being the storyline projected on the plane of graphs which are representing the networks of characters. To keep the storyline, a possible solution was proposed in [3]; in this reference we have shown how to build some time series, and, by means of them, analyse the literary work in the same manner as if we were processing a sequence of data measured at successive times. In this paper, we will continue the discussion on time series, applying them to a poem, adding also some examples of the use of graphs [8].

Actually, the digitized form of a literary work can be used to build some sets of numerical data which can be analysed by means of algorithms and mathematical models [3]. For instance, we can collect the occurrences of any word in the text and evaluate its relevance. From this analysis, we gain an appraisal of the importance of characters, places or leitmotivs of the literary work, to which the selected word is linked. The data of the word occurrences can also be arranged as a time series, where 'time' is played by the sequence of lines of the digitized text. In this manner, the word occurrence becomes a function of time. As we are doing for functions, we can plot the series to have a visual perception of it. The plot, for instance, can be made in the manner proposed in Ref.3, which is giving a barcode-like image.

In Reference 3, we discussed the use of time series and co-occurrence matrices in the analysis of some characters of the novel Harry Potter and the Philosopher's Stone. Here we show examples of time series obtained from Dante Alighieri's poem, the Divina Commedia. Instead of searching for the role of characters, we use time series and graphs to illustrate some leitmotivs of the Comedy. We will be able to easily verify that two of them are 'Love' and 'Light', and that they are overwhelming terms in the Paradise. This is in agreement with the view of the Comedy as a splendid illustration of Dante's metaphysics of Divine Love [9]. Before discussing time series and graph, let us spend some words on this poem.

2. The Divine Comedy

The Divine Comedy, Divina Commedia, is the poem written by Dante Alighieri, from approximately 1308 and his death in 1321. This poem is an allegorical journey in the realms of afterlife, based on the vision of the world of medieval Western Europe [10]. In the poem written in the first person, Dante describes his journey through Hell and Purgatory, and then his rise through Heavens up to the Empyrean and the vision of God. This travel is representing allegorically a soul's journey towards the Divine Love of God.

The Comedy is composed of 14,233 lines, divided into three Cantiche: Inferno (Hell), Purgatorio (Purgatory), and Paradiso (Paradise). Each is consisting of 33 canti. The total number of canti is 100, since there is an initial canto, which serves as a prologue to the poem. In the first canto, when Dante is in a dark wood, he meets the Roman poet Virgilio. In the third canto, Dante and Virgilio enter the Hell. Virgilio guides Dante through Hell and Purgatory;

✉ d002040@polito.it, amelia.sparavigna@polito.it

Beatrice, Dante's ideal woman, is leading him through the Paradise.

The structure of the three realms is given in a numerical pattern of 9 plus 1, for a total of 10. There are 9 circles in the Inferno, and Lucifer is contained at its bottom, at the center of the Earth. We have 9 rings on the Mount Purgatory, and the Garden of Eden is at its top. Then, Dante finds the 9 'cieli', the celestial spheres of Paradise crowned with the Empyrean, and God, the radiating point of Light. In each of the circles, rings and spheres, Dante and his guides meet different characters, because each of these places is concerning a different sin or virtue.

3. Word occurrences

For machine-readable texts several mathematical models exist, useful to study the style of literary works, for instance, by the frequency of word counts, vocabulary items, or grammatical forms (references are given in [3]). Here, the analysis focuses on the use of time series, as we did in [3]. We will investigate the occurrence of words and, as a consequence, of some leitmotivs of the poem in the canti composing the Divine Comedy. In fact, we consider the leitmotivs playing the role of 'characters' and the canti that of 'places'. For instance, the first canto of Inferno is the 'place' of a dark wood, and in it we find the 'character' of Fear, but also those of 'Love' and 'Hope'.

Aim of this paper is not a detailed analysis of the Comedy, but that of showing some examples of a method based on time series. Then, we choose just a few words: Love (Amor or Amore), Light (Luce, Lume), Desire (Disio), Hope (Speranza), and someone else. Let us start from the word Love: we can search the lines where 'Amor' or 'Amore' are occurring. In this manner, we have a time series where the line number of the poem is representing the time; the value of the series is 1 when the word is occurring and 0 when it is not occurring. After we have obtained the time series of Love, we can plot it: the time is on the horizontal axis, and when the series has value 1, a vertical line is plotted. Once the series is totally plotted, the result is a barcode-like image representing it. In the Figure 1, some time series are shown. From their barcodes, we can easily appreciate the dominant motifs of the three realms: Paradise is triumph of Light (Luce, Lume) and Divine Love. Fear, but also Dante's Hope of saving his soul, are the dominant leitmotivs in the Inferno.

4. Graphs of canti and words

In [2] and [7], we have proposed for the analysis of literary works the use of Graphviz, the Graph Visualization Software, a package of open-source tools for drawing graphs specified in DOT language scripts. Graphviz is a free software under Eclipse Public License, available at the Web address http://graphviz.org/. This software has a quite simple description language; moreover, we can prepare graphs with several architectures, for instance with the "circo" layout, to produce a circular plot. In [2], we used Graphviz in the analysis of Shakespeare's Hamlet and Othello, to display the networks of characters.

As previously told, in this preliminary work on the Divine Comedy, we prefer analysing not characters but the dominant motifs of the poem. Therefore, we use Graphviz to represent the occurrence of some words in a place (canto). Here then, the proposed graphs have words and canti as vertices, and edges that are giving the occurrence of a word in a canto. This means the following: if we see an edge linking a word to a canto, it means that this word is occurring in it. The Figure 2 is showing an example: it is giving the occurrence of some words (Amore, Paura, Lume, Notte, Pena, Luce, Dolore, Fede, Disio, Giustizia, Speranza, Morte) in the first five canti of Inferno. For the sake of simplicity, we use just a few words and canti.

The graph was made using Graphviz with 'circo' filter to have a circular layout. From the graph in the Figure 2, we can see that all the words of the list are appearing at least in one of the canti. To have the Figure 3, concerning the first five canti of Purgatorio, we used the same list of words (Amore, Paura, Lume, Notte, Pena, Luce, Dolore, Fede, Disio, Giustizia, Speranza, Morte). We can see that some of these words are not occurring in these canti of Purgatorio. They are Giustizia, Dolore and Fede. In the Figure 4, we considered the same list of words for the first five canti of Paradiso. Let us note that a lot of words of the list are not occurring there. These words are Pena (Punishment), Paura (Fear), Notte (Night), Dolore (Pain), Morte (Death), Speranza (Hope). This is not surprising: Dante is in the Paradise, and then his soul is free from these motifs. Light and Love are dominating his spirit as all the Paradise. The proposed example is then showing how graphs can illuminate the leitmotivs of a poem. In the Figure 5, we are showing again the graph given in Figure 4, but in this image, near the edges of the graph, we have added the number of word occurrences in each canto. For instance, number 3 near the edge linking Disio to Canto IV, is indicating that this word is occurring three times in this canto.

5. Discussion

In the given Figures, we have shown the results of an occurrence analysis just for a few words and canti. Of course, this approach can be used with much more terms. It could be proposed for characters too, inside each canto, with edges linking characters and leitmotivs. Let us remember that during his journey,

Dante and his guides meet several souls, which are different in each canto. Of course, this is a consequence of the structure of the realms of the afterlife. Then, the analysis on characters and leitmotivs is significant when it is made specifically for each canto. For what concerns an analysis with CHAPLIN, the software proposed in [7], it is possible too and under development in future works.

Let us conclude this paper with a note on two of the words occurring in the Comedy, Lume and Luce. Dante used two terms for Light, terms that are coming from two Latin words, Lux and Lumen. Lux is the clear, bright light, or the light of the day. It is a term derived from the Proto- Indo-European root leuk- (bright, white light) [11]. From this root, the Greek 'Leuko' for White is coming. From Lux, the Latin derived the word Lumen (light, lamp, torch, source of light), and the derivations luminare, luminosus, illuminare. The Italian has Lume in a figurative meaning too: "lume della ragione", that is, "light of reason". It means that 'Lume' can be used to indicate a source of light, or a light received by enlightenment. Let us give two examples from the Comedy. Dante is using Lume when he is meeting Virgilio, and his invocation starts with "O de li altri poeti onore e lume", and when Dante is looking at Beatrice, he sees the "lume de la dolce guida". Lume is also used to define some souls of the Paradise. 'Luce' is used for the Light of God, the 'luce etterna'. Dante is therefore maintaining in Italian the use of the medieval Latin of splitting 'light' into two elements, Lux and Lumen. We have seen this splitting clearly shown in a text of Robert Grosseteste on the metaphysics of light and creation of the world [12].

References

1. A.C. Sparavigna, On Social Networks in Plays and Novels, International Journal of Sciences, 2013, Volume 2, Issue 10, Pages 20-25.
2. A.C. Sparavigna and R. Marazzato, Graph Visualization Software for Networks of Characters in Plays, International Journal of Sciences, 2014, Volume 3, Issue 2, Pages 69-79.
3. A.C. Sparavigna, Co-occurrence Matrices of Time Series Applied to Literary Works, International Journal of Sciences, 2014, Volume 3, Issue 7, Pages 12-18.
4. F. Moretti, Network Theory, Plot Analysis, A Stanford Lit Lab Pamphlet, 2011, litlab.stanford.edu/ LiteraryLabPamphlet2A.Text.pdf
5. J. Stiller, D. Nettle and R.I.M. Dunbar, The Small World of Shakespeare's Plays, Human Nature, Volume 14, Issue 4, 2003, Pages 397-408.
6. R. Alberich, J. Miro-Julia and F. Rossello, Marvel Universe Looks Almost Like a Real Social Network, arxiv.org/abs/ cond-mat/0202174v1.
7. R. Marazzato and A.C. Sparavigna, Extracting Networks of Characters and Places from Written Works with CHAPLIN, arXiv:1402.4259 [cs.CY], 2014.
8. Oystein Ore, Graphs and Their Uses, New Mathematical Library Series, 1990, Volume 34, Cambridge University Press.
9. A.J. Smith, The Metaphysics of Love: Studies in Renaissance Love Poetry from Dante to Milton, Cambridge University Press, 1985.
10. A.C. Sparavigna, From Rome to the Antipodes: The Medieval Form of the World, International Journal of Literature and Arts, 2013, Volume 1, Issue 2, Pages 16-25.
11. J. P. Mallory and D.Q. Adams, Encyclopedia of Indo-European Culture, 1997, Taylor & Francis, Page 513.
12. A.C. Sparavigna, Robert Grosseteste's Thought on Light and Form of the World, International Journal of Sciences, 2014, Volume 3, Issue 4, Pages 54-62.

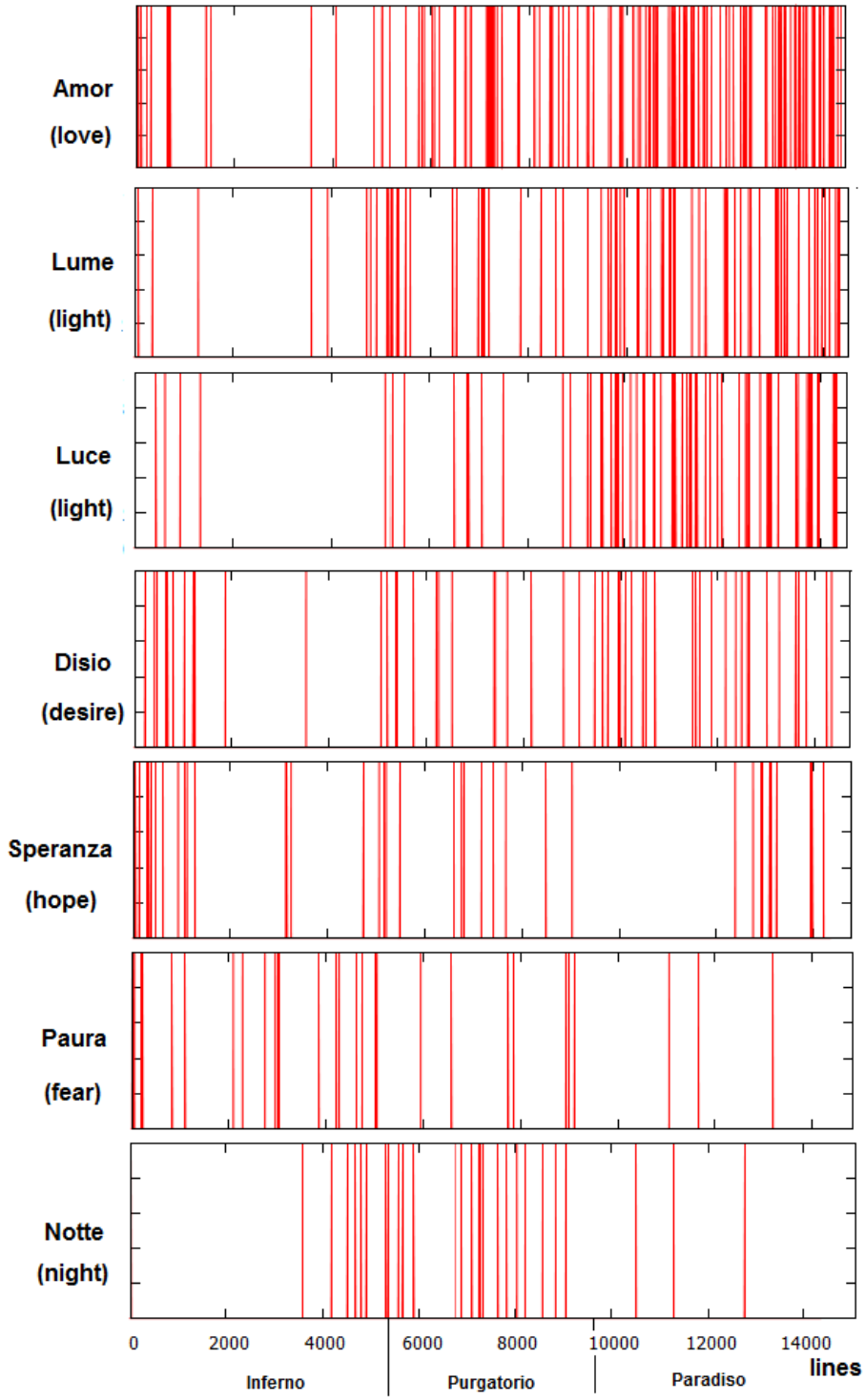

Figure 1 – Time series representing the occurrence of some words in the Divine Comedy. On the horizontal axis, line numbers of the poem are representing time. Let us stress that the poem is subdivided in three parts (Hell, Purgatory and Paradise). The vertical red lines represent the word occurrences; consequently, the plot produces a barcode-like image. From these barcodes, we can easily appreciate the dominant leitmotivs of the three realms: Paradise is triumph of Light (Luce, Lume) and Divine Love.

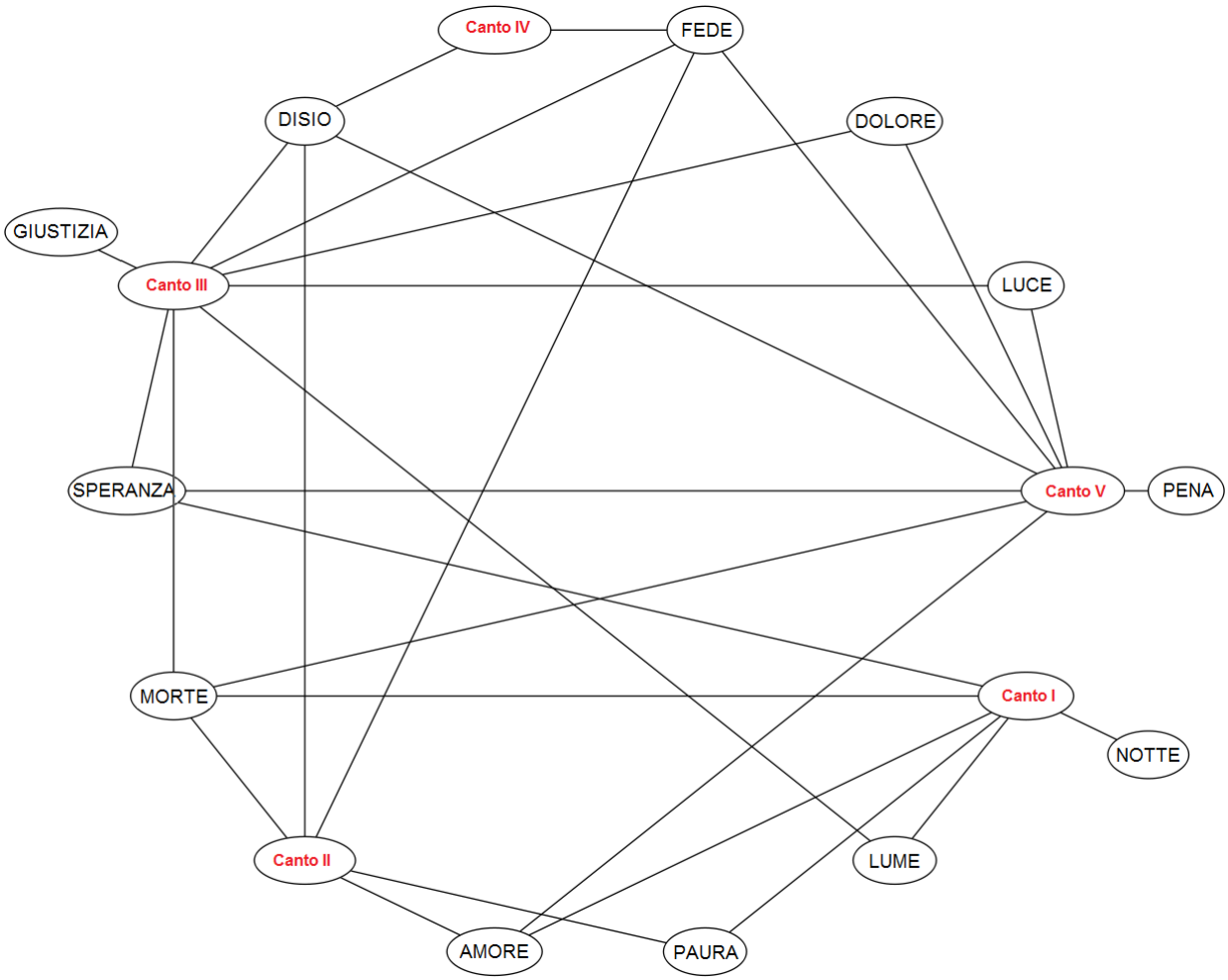

Figure 2 - This graph is showing the occurrence of some words (Amore, Paura, Lume, Notte, Pena, Luce, Dolore, Fede, Disio, Giustizia, Speranza, Morte) in the first five canti of Inferno. The graph was made by Graphviz, applying the 'circo' filter to have a circular layout. The vertices of the graph are words and places (canti). Edges are linking words and canti, and therefore are representing the occurrence of a word in a canto. We see that all the words of the given list are linked to a canto. Disio (Desire), for instance, is occurring in Canti II, III, IV and V, but not in the first canto, as it is shown by the four edges. Amore (Love) is in Canti I, II and V; Giustizia (Justice) only in the third canto. This word is in the inscription on the Gate of Hell: "Justice the founder of my fabric moved".

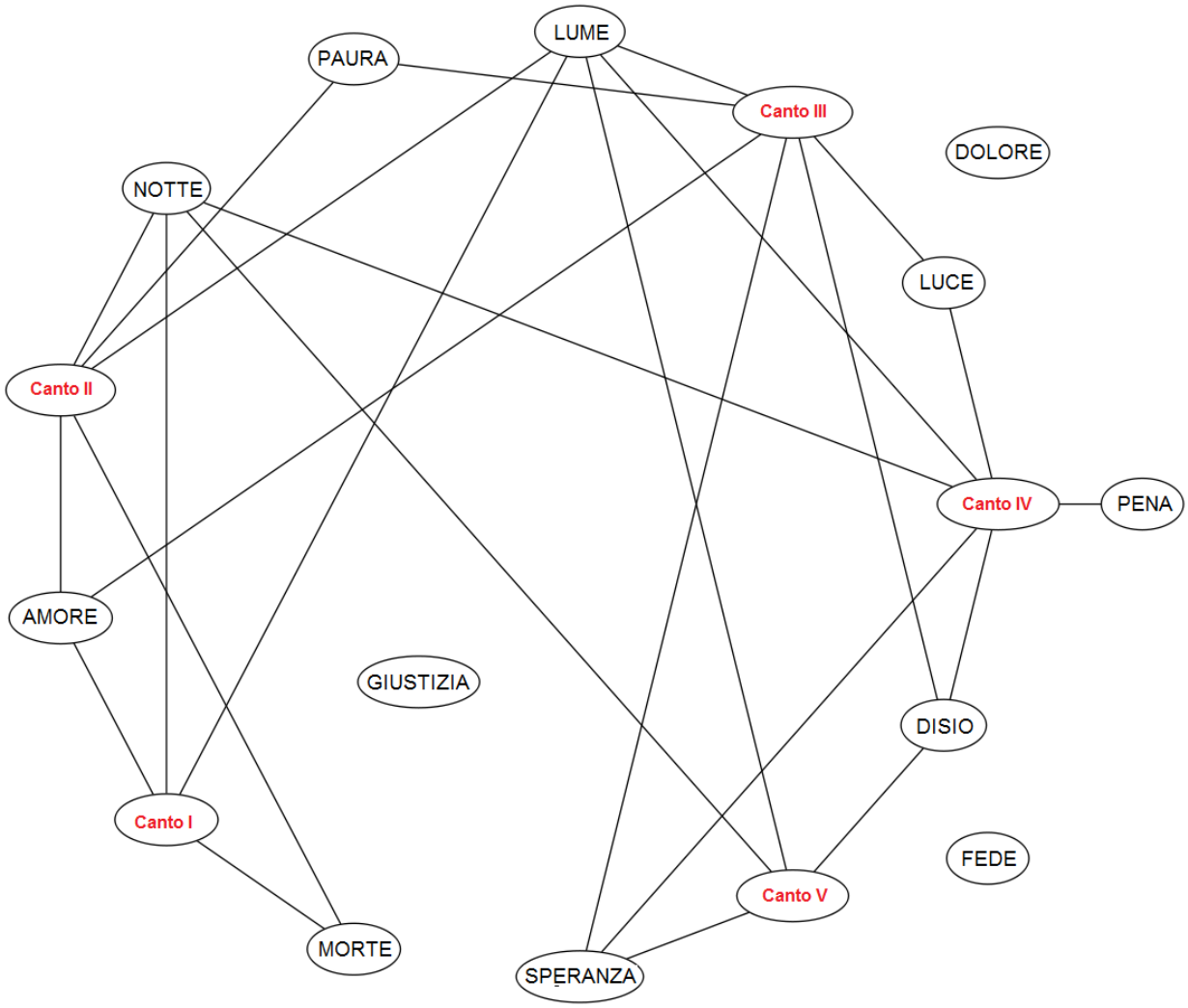

Figure 3 – This graph is showing the occurrence in the first five canti of Purgatorio of the words (Amore, Paura, Lume, Notte, Pena, Luce, Dolore, Fede, Disio, Giustizia, Speranza, Morte), that we used for Figure 2. The vertices of the graph are words and places (canti). The edges are representing the occurrence of a word in a canto. Let us note that some of the words are not occurring in these canti of Purgatorio, and therefore are shown as isolated vertices. They are Giustizia, Dolore and Fede. Disio (Desire) is in Canti III, IV and V. Note that Lume (Light) is in all the five canti. The leitmotivs of Purgatorio are different from those of Inferno.

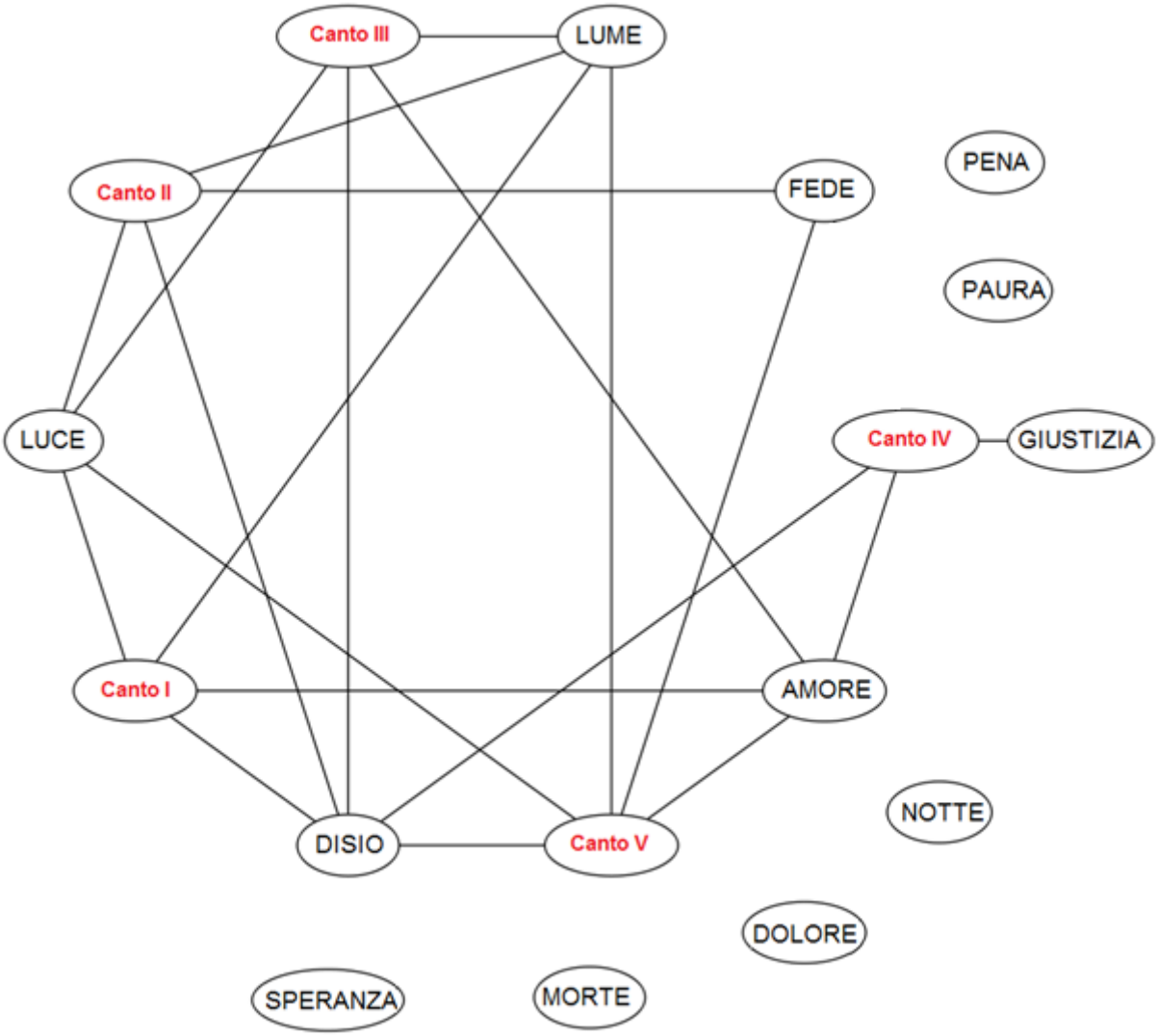

Figure 4 – This graph is showing the occurrence in the first five canti of Paradiso of the same words used for Figures 2 and 3. Let us note that many words are not occurring in these canti, and therefore are represented as isolated vertices. These words are Pena (Punishment), Paura (Fear), Notte (Night), Dolore (Pain), Morte (Death) and Speranza (Hope). This is not surprising: Dante is in the Paradise, and then his soul is free from these motifs. Moreover, Light and Love are dominating in Paradise, as we have seen in the Figure 1. The proposed example is then showing how graphs can illuminate the leitmotivs of a poem.

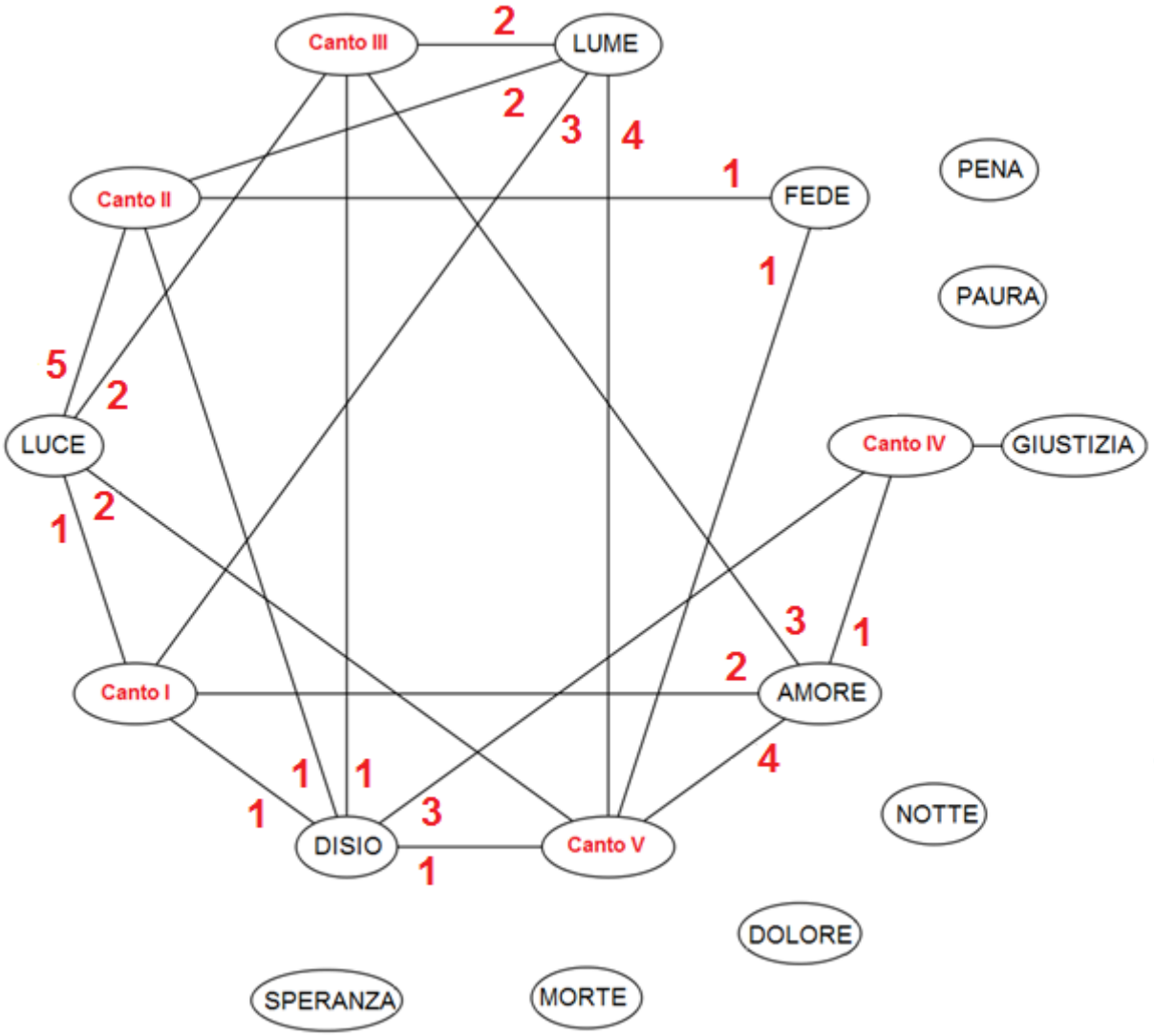

Figure 5 – As in the previous Figure, this graph is showing the occurrence in the first five canti of Paradiso of the words Amore, Paura, Lume, Notte, Pena, Luce, Dolore, Fede, Disio, Giustizia, Speranza and Morte. However, in this image, near the edges of the graph, we have added the number of occurrences of words in each canto. For instance, number 3 near the edge linking Disio to Canto IV, is indicating that this word is occurring three times in the canto.