Exploiting networks in Law

Guido Boella, Luigi Di Caro, Andrea Violato, Livio Robaldo
Department of Computer Science, University of Turin,
Nomotika s.r.l. (http://www.nomotika.it)
{boella, dicaro, robaldo}@unito.it, violato.a@gmail.com

Abstract

In this paper we first introduce the working context related to the understanding of an heterogeneous network of references contained in the Italian regulatory framework. We then present an extended analysis of a large network of laws, providing several types of analytical evaluation that can be used within a legal management system for understanding the data through summarization, visualization, and browsing. In the legal domain, yet several tasks are strictly supervised by humans, with strong consumption of time and energy that would dramatically drop with the help of automatic or semi-automatic supporting tools. We overview different techniques and methodologies explaining how they can be helpful in actual scenarios.

Keywords: Networks in Law, Network Analysis, Legal Management Tools

1. Introduction

The importance of Information Technology tools lies on the ability of limiting the well-known complexity of the regulatory framework. Starting from the language used to express concepts and rules that often create serious problems of interpretation, *scientia iuris* is full of innumerable aspects that make it difficult to manage by domain experts. The use of computer systems can reduce this complexity for those who have to do with a legal reality that results to be more and more interdisciplinary, international and multifunctional. All this is even more true by taking into account one of the most complex aspects of any modern legal text, which is the extensive use of *references*. It is therefore the intention of this paper to first investigate the need for the development of tools capable of assisting the employee to manage the particular regulatory complexity characterized by cross-references, and then to simulate possible technological solutions.

In (Boella et al., 2011) we introduced the Eunomos software, which is being developed in the context of the ICT4LAW project[1]. Eunomos is an advanced legal document management system based on legislative XML representation of laws which are retrieved automatically from institutional legislative portals, and incorporates a tool for building legal ontologies called Legal Taxonomy Syllabus (Ajani et al., 2008). It provides a powerful knowledge base for analysing laws and keeping up to date with legal changes, although it also requires knowledge engineers to manage the information. Eunomos can be regarded as a desktop where several tools are at the disposal of the legal knowledge engineer to assist with categorising and annotating laws and building ontologies. In this work, we use the data provided by Eunomos to perform network analysis experiments on its large network of references between laws.

[1]ICT4LAW: ICT Converging on Law: Next Generation Services for Citizens, Enterprises, Public Administration and Policymakers funded by Regione Piemonte 2008-2013, call Converging Technologies 2007, website: http://www.ict4law.org

2. Uses of Normative References

In this section we present an overview of actual scenarios in which references play crucial roles for an accurate understanding of the regulatory framework, and where automatic tools can represent a matchless support.

2.1. Completion

The presence of thousands of stratifications in the Italian legal sources over the years has enabled the strengthening of the use of cross-references between legal texts in order to complete the content. In this sense, emblematic is the discipline offered by *d. lgs. June 8, 2001, no. 231*, which introduced the administrative liability of legal persons as a result of the commission of offenses having penal account[2]. The *cc.dd. "Underlying Offenses" (Reato Presupposto) ex Articles 24 et seq.* (which, if made by individuals, apical or subordinates, in the interest of the institution or to the advantage of the same, it allows the attribution of administrative liability for the institution itself) are presented as an ordered and constantly-changing list of references to other texts, each of which represents the logical handhold to which entrust to really understand the regulatory scope of the decree[3]. For instance, let us consider the first paragraph of Article *24*:

> 1. In relazione alla commissione dei delitti di cui agli articoli 316-bis, 316-ter, 640, comma 2, n. 1, 640-bis, 640-ter se commesso in danno dello Stato o di altro ente pubblico, del codice penale, si applica all'ente la sanzione pecuniaria fino a cinquecento quote.
>
> *[1. In relation to the commission of the crimes referred to in Articles 316-bis, 316-ter, 640, paragraph 2, no. 1, 640-bis, 640-ter if committed*

[2]In the field, innovative in many ways in 2001, in addition to reverse one of the key concepts of the legal Italian thought (brilliantly enclosed in the Latin *Brocardo*, now abandoned, *societas delinquere non protest*) it presents a structure that allows us to clearly understand the complexity of the references in a legal text to provide the completion of the content of the rules.

[3]In the paper, we will use the single term "decree" to intend "law decree".

against the State or other public body of the Criminal Code, apply monetary sanctions of up to five hundred shares.]

The resulting framework is very complex, because the reconstruction accomplished by the normative scope of *d. lgs. 231/2001* neccessarily needs a coordinated and parallel reading of several elements of the criminal code, which inevitably brings to a costly consolidation of different texts. Automatic tools for retrieving and visualizing missing information can speed up the process.

2.2. Intervention

This scenario is even more complex when considering a second and unique aspect of this discussion: the temporal dynamism that characterizes the use of normative references in any legal context. In fact, it is not a rarity that, over the years, the legislature put hand to the subjects adapting the contents to the stimuli given by socio-economic factors and the surrounding cultural context. This natural evolution is often mediated through a surgical use of references which permit to individuate the precise points under modification. In the Italian scenario, this assumption is far from being uncommon. Many, in fact, are the examples of legal decrees subsequently converted (even with modifications) into laws.

For instance, the decree of the 14 August 2013, *n. 93, Article 9* has resulted in the introduction of the third paragraph of Article *640-ter* of the Criminal Code (Replacement of digital identity) and the amendment to Article *24-bis ex d. lgs. 231/2001* (Computer crimes and illegal data processing) with the provision of new crimes on digital identity, money laundering and protection of privacy. Let us consider the following example.

> *1. Article 640-ter of the Penal Code, shall be amended as follows:*
> *a) after the second paragraph, the following is inserted: "The punishment of imprisonment from two to six years and a fine from EUR 600 to EUR 3,000 if the act is committed by replacement of identity to the detriment of a one or more subjects."*
> *b) the last paragraph, after the words "in accordance with" the following is inserted: "and the third".*
>
> *2. Article 24-bis, paragraph 1 of the Decree of 8 June 2001, n. 231, the words "and 635-d" are replaced by the following: "635-d and 640-ter, paragraph," and after the words "penal code" the following is added: "and the crimes of which articles 55, paragraph 9, of Legislative Decree 21 November 2007, n. 231, and in Part III, Title III, Chapter II of legislative Decree 30 June 2003, n. 196."*

The example reported is useful to understand how normative references act, both in terms of intervention on amending a discipline and of completion of content.

Indeed, it is through the use of references that was made possible an intervention amending the penal code and the text of the *d. lgs. 231/2001*, adding to the provisions already present in both sets of rules, new cases. Still, it is always through the use of normative references that, in the specific legal reality designed by *d. lgs. 231/2001*, disciplinary framework is completed[4].

2.3. Understanding

Finally, it is worth noting how the use of normative references can also take place outside the purely legislative context. In fact, it is quite common to find a citation of a law inside a legal text that is not used with explicit references, but also by case law of the courts, which, in exercising a peculiar hermeneutical activity of this rule, it implicitly complements the content, clarifying the more conceptual and obscure aspects and thus bringing important suggestions to the reader.

A useful example to understand this particular use of the standard references in the Italian world of law is what happened in the field of copyright with the important conceptual clarification provided by the Supreme Court of *Cassazione* (Criminal Section 3), with the judgment of 22 December 2009, *n. 49385*, in which it is explained that:

> *The offense provided by Art. 171a, first paragraph, first sentence, second hypothesis, of the Law of 22 April 1941 no. 633 (illegal possession, for commercial purposes or Entrepreneurial, computer programs devoid of mark SIAE[5]), when it requires that detention is for "commercial or business purposes", it does not refer also to the detention and use as part of a professional activity, to which therefore the norm in question does not apply.*

This information are, to the expert, essential to the understanding of the rule, since they are potentially difficult to acquire without a meticulous connection between legislation and its judicial clarification.

3. Technologic Background and Related Work

In this section we overview the fields and the technology that can represent the core of automatic tools supporting the forementioned uses of the normative references.

References among laws can be seen as links in a graph. Graph Theory is the field that studies the properties of a specific model of data, i.e., relations between objects. Generally speaking, a graph is in fact made up of nodes and links that connect them. If the links have no direction, the graph is called *undirected*, otherwise we talk about *directed graphs*. (Bollobás, 1998) presents a large overview over the

[4] As a decree, for which it may be enacted only for emergencies and in cases of particular gravity by the Council of Ministers, and that has a 60-days validity after its enactment, it is not possible to know if and when will be its conversion (with possible modifications) in ordinary law of the state.

[5] Societa' Italiana degli Autori ed Editori [*Italian Society of Authors and Publishers*].

mathematical properties of graphs. In our case, when a law changes or cites another law, the opposite may not subsist, thus we are in the case of directed graphs.
Graphs are usually not regular in real systems. This means that there is no perfect order or disorder, and some nodes result to be more connected than other ones. Conversely, in a random (i.e., artificial and homogeneus) network, the probability of having a link between a pair of nodes is equal for all possible pairs.
Several datasets from different domains can be modeled through a network, and there is a large literature on network analysis techniques that permit to extract relevant information or to support navigation in general. While most of the efforts has been put on Social Networks, there are several other applications in other domains. To the best of our knowledge, in the legal domain there are still few attempts that leverage such methods to create methodologies and applications to help domain experts in finding and understanding legal data in real scenarios.
For example, in (DE et al., 2006) the authors presented a parser to resolve references in legal text. However, they do not provide any methodology nor tool for an active analysis of the data. In a similar way, (Keune et al., 2012) used the network contained within a single article in the Dutch law to extract and weight internal and external references, visualizing them in a graph. Finally, (Winkels and de Ruyter, 2012) investigated the role of citations in the case law of the Supreme Court of The Netherlands for identifying semantic information concerning their authority.

4. Techniques and Applications

In legislation, all the data can be also viewed as complex networks where nodes are laws and links represent kinds of relationship like "*modification*", "*implementation*", "*substitution*", and so on. These information are often complex to treat, organize, and use since they are many and continuosly changing over time. Manual intervention, indeed, is often required in classical tasks because of the difficulty to get the "big picture", that is, to have an at-a-glance overview over the data under evaluation.
Social Network Analysis is quite a new field that inherits methods from Phisics (i.e., Complex Networks) and Mathematics (i.e., Theory of Graphs) to face problems related to the huge amount of data coming from social networks like Twitter[6], Facebook[7], Flickr[8], and so forth. These algorithms are useful to capture statistics about the type of the networks along their properties. More in detail, it is possible to analyze a network in terms of its evolution over time, important nodes, implicit relationships, etc. All these evaluations can be helpful also in a network of laws. In this section we present a set of analyses that may help the jurist to deal with these networks.

4.1. Communities

The first statistical analysis of a graph is given by the distribution of the number of neighbours of a node, also called *degree*. In real networks, the degree distribution has a tail that often follows a power law, that means that it contains many nodes with low degree and some node with large degree. This characteristic of real networks is called *community structure*. Such communities are also commonly named clusters, and they represent nodes that play similar roles within the graph. In Figure 1 is shown a subset of the graph used in our evaluation (around 7.0% of the total dataset) where it results clear the different degree of the nodes, since the used visuaization strategy maps the most connected nodes towards the center of the screen while the low-degree nodes lie on the borders.

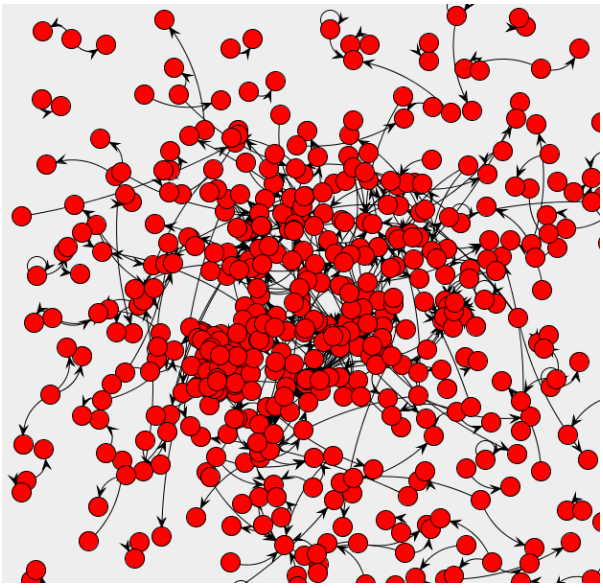

Figure 1: A 500-links subset of the graph of laws used in the analysis where it is visible the natural heterogeneity of the nodes in terms of *degree*.

A common and often useful analysis of a network looks at such latent community structures, and it is based on some clustering approach. In our test we made use of the concept of *edge betweenness* of an edge, i.e., the extent to which it lies along shortest paths between all pairs of nodes in the network. This algorithm works by iteratively following two steps: computing the edge betweenness for all edges in the current graph and then removing the edge with the highest betweenness value. This analysis helps finding communities, since it iteratively removes central nodes to separate the graph in distinct subnetworks.

4.2. Important Nodes

Relationships between nodes of a network may have a precise direction, that needs to be taken into account to understand the system as a whole. PageRank (Page et al., 1999) is an algorithm that assigns a weight to each node of a network in the World Wide Web domain, with the purpose of quantifying its relative importance within it. In spite of its original use on hyperlinks, it can be useful for several other domains dealing with directed graphs. Generally speaking, in the legal domain, if one law modifies a law B, one usually does not find on B a link back to A. In some case, there can be few relationships that may be reciprocal (i.e., cita-

[6] http://www.twitter.com/
[7] http://www.facebook.com/
[8] http://www.flickr.com/

tions). The PageRank algorithm measures the importance of a node by considering that of the nodes that link to it. Thus, this indicates a finer node evaluation with respect to the edge betweenness. An accurate analysis of the most important nodes within a network of laws can be helpful to understand and use the data.

4.3. Diameter Analysis

The *diameter* of a graph can give useful insights. In detail, it is the largest number of nodes which must be traversed in order to travel from one node to another. In a network of laws it can be used to estimate the maximum complexity of modification/citation paths. Since a network of law is often disconnected, it is constituted by many subnetworks. In this case, the diameter considers the maximum distance found in all subnetworks.

5. Data and Results

In this section we present an analysis of the data coming from Eunomos, i.e., a network of around 10K laws interconnected by 7K links. The dataset contains different types of laws, as for instance "*stato:legge*", that indicates the principal type of law of Italy. Table 1 illustrates all the existing types along with their meaning and cardinality.
On the other hand, in addition to the law types, there exist different kinds of relationship. For instance, the most frequent link type contained in the data is "*abrogation*", that indicates that a law *A* removes the validity of another law *B*. The complete set of link types is shown in Table 2.

Type of law	Cardinality
Modification	2456
Abrogation	3809
Introduction	821
Others	16
Total	7102

Table 2: Types of relationship and relative cardinality.

In Figure 2 is then shown the distribution of the degree levels. As one can see, most of the laws have few connections with other laws, whereas some other laws have a large connectivity within the graph. This is in line with the majority of the networks automatic systems have to deal with.
The diameter of the network (actually, the maximum diameter of all the disconnected subnetworks in it) is 8. This means that, in the worst case, if a jurist has to navigate and understand a path between two laws in the network, it could go through other eight laws. Still, links can have different meanings (see Table 2), so the process may include multiple cateogories of operations. Having a tool that can support this navigation may represent a dramatically help in such process.
Finally, Figure 3 shows the distribution of the PageRank scores over the nodes/laws of the network. Notice that this analysis gives an interesting result since it seems to perfectly split the important nodes (also called hubs) from the others. A system that makes use of these scores can filter out unimportant information rather than letting emerge crucial points within the regulatory framework.

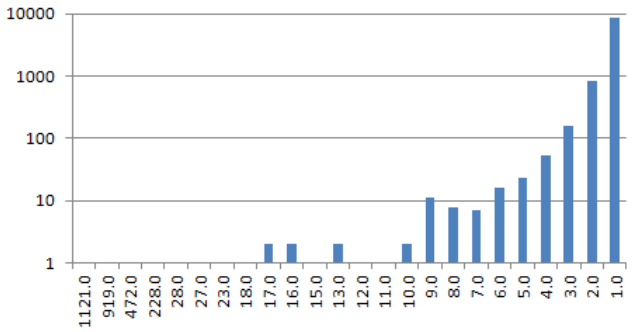

Figure 2: The degree distribution of the network extracted from the dataset (note that we used a logaritmic scale). Most of the laws are linked to few other laws (long tail). From the other side, two laws have around one thousand connections (1121 and 919 respectively).

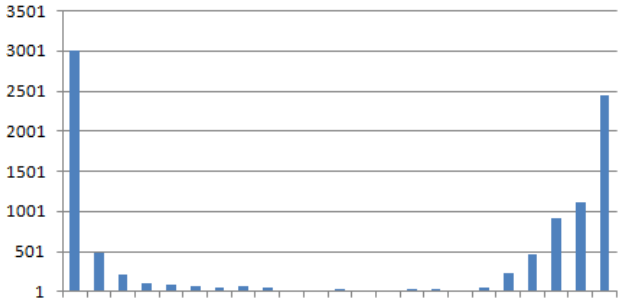

Figure 3: The distribution of the PageRank values of the network extracted from the dataset. Notice that this analysis gives an interesting result since it seems to perfectly split the important nodes (i.e., hubs) from the others.

6. Conclusions

Information Technology can be an important collaborator for one who is daily struggling with the consolidation of a fragmented and heterogeneous regulatory framework that is often difficult to identify in a rose of precise factors. By partially automating the completion of the content of a legal rule, or through suggestions of references to the case law, it is possible to optimize the productivity of the expert which can be relieved of the problems arising from the activities of reconstruction and interpretation.
In this work we presented a dataset containing a network of Italian laws, carrying an extensive analysis through several types of analytical techniques that can be used within a legal management tool for understanding the data through summarization, visualization, and browsing. In the legal domain, yet several tasks are strictly supervised by humans, with strong consumption of time and energy that would dramatically drop with even little help from automatic systems. We overview different techniques and methodologies explaining how they can be helpful in classical scenarios.

Type of law	Cardinality	Meaning
stato:legge	4050	The ordinary law of the State is the source of law for excellence and is produced by Palamento (the legislative body). It is used to legislate on any domains.
presidente.repubblica:decreto	1452	Although its function is to represent the country, the President of the Republic retains its specific power to legislate in certain cases and domains (government regulations, appointment of the chairman of the board of ministers and individual ministers, appointment of the chairmen of the House and Senate)
stato:decreto.legge	1108	It is an act of emergency issued in case of special need and with a limited validity time (60 days, after which Parliament transforms and incorporates them by ordinary law or it decades)
stato:decreto.legislativo	840	It is an act of the executive power of a specific parliamentary mandate (through a framework law, which sets out the powers and the issues of the decree). Once issued, it is comparable to an ordinary law of the Parliament (it basically changes the name).
presidente.consiglio.ministri:decreto	65	The ministerial decree is a source of law in all respects, though emanating from a specific Ministry, signed by the President of the Council of Ministers.
stato:regio.decreto.legislativo	15	It is an act of the monarchic era. There are still few regulations and they will be incorporated in future disciplines.
Others	2027	-
Total	9557	

Table 1: Types of laws contained in the dataset along with their cardinality and meaning.

In future works, we aim at deepening the usefulness of such technology by implementing practical tools and by carrying extensive studies with domain experts.

7. Acknowledgements

The work has been funded by project ITxLaw financed by Compagnia di San Paolo.

8. References

Ajani, G., Boella, G., Lesmo, L., Mazzei, A., Radicioni, D. P., and Rossi, P. (2008). Legal taxonomy syllabus: Handling multilevel legal ontologies. *Proceedings of Langtech 2008*.

Boella, G., Humphreys, L., Martin, M., Rossi, P., and van der Torre, L. (2011). Eunomos, a legal document management system based on legislative XML and ontologies (position paper). *Applying Human Language Technology to the Law*, page 41.

Bollobás, B. (1998). *Modern graph theory*, volume 184. Springer.

DE, E., Winkels, R., and van Engers, T. (2006). Automated detection of reference structures in law. *Frontiers in Artificial Intelligence and Applications*, page 41.

Keune, A., Intelligentie, B. O. K., Winkels, R., Boer, A., and BG13a, B. (2012). Finding relevant information in dutch legislation with network analysis.

Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: bringing order to the web.

Winkels, R. and de Ruyter, J. (2012). Survival of the fittest: network analysis of dutch supreme court cases. In *AI Approaches to the Complexity of Legal Systems. Models and Ethical Challenges for Legal Systems, Legal Language and Legal Ontologies, Argumentation and Software Agents*, pages 106–115. Springer.