

Towards a Measuring Device of Legal Complexity

ACLE LAW research seminar, Amsterdam 27 May 2015

Marijn Koolen and Bart Karstens

We present the first steps to come to a measuring device of the complexity of legal texts. The datasets we currently work on include national constitutions and civil codes. However, the device we develop should also be applicable to other legal documents. A measure of legal complexity is given by linguistic and structural characteristics. With respect to linguistic characteristics we measure such things as length of text, number of words, number of unique words and Flesch reading ease. With respect to structural properties there are two things to consider. First, there is the general structure of the text in terms of levels and sublevels (chapters, articles, sections, etc.). Second, there are the references in the text to other parts of the same text, or to other texts. This is the most innovative part of our research. Conceiving of a legal text as a network of relations, enabling analysis of legal documents through concepts stemming from social network theory, is a fresh line of research in which there is ample room for investigation. Most of the research that is available has focused on how the law functions as a system of interdependent texts, in accordance with methods from citation analysis. Our aim is different. We use network analysis in order to gain deeper insight into the structure and complexity of *singular* legal texts. With this goal in mind we measure network properties such as indegree, outdegree, closeness, betweenness, eigenvector centrality and various clustering coefficients. When all indicators of linguistic and structural complexity are taken together it is possible to come to a statement of the complexity of legal documents. This promises to clarify a number of things. First, we want to gain deeper insight in the nature of legal complexity. What is a metric statement of 'normal' complexity? How does such a statement relate to typical networks? Should we aim to reduce complexity as much as possible, as a number of scholars have suggested? What causes complexity? Can we infer the process of construction of a text from the network properties of the text? To answer these questions we use cost-benefit and affiliation mode analysis. Second, we offer a new perspective on comparative law. Our device allows for classification of countries where legal family ties are based on metric properties. Finally, we are interested in dynamics *on* the network. With the network representation we can assess the efficacy of the law, both in terms of enforcement of the law itself and in terms of the law as a mechanism for social and economic organization. We include a paper on the network properties of national constitutions in which we present a number of interesting results. As it turns out not all forms of referential complexity have a negative effect on readability of constitutions. On the contrary, particular reference structures enhance interpretability of these legal documents. The presence of hubs and authorities, for example, or a low cluster component ratio, sort this positive effect. Further, we gained insight in the functionality of the central nodes. This can be used to set up a typology of internal references. Finally, a very clear distinction between civil law (few internal references) and common law countries (extensive referencing) presented itself. In our presentation we expand these results with recent experimental findings. We also discuss the key problems we encountered in the construction of our parser to retrieve the reference structures and the choices we have correspondingly made.