

Legal Sources and Hyperstructures

Kirill Miazine
NRCCL, University of Oslo
kirill.miazine@student.jus.uio.no

01.05.2006

This document aims to provide a short summary of my student thesis written while I was research assistant at NRCCL.

The basic idea and the inspiration source for my work was that legal documents are massively interconnected. There are two different kinds of connections that can be identified: syntactic and semantic.

Current systems for legal information retrieval (LIR) support searching for relevant documents using free-text search. The search facility is commonly based on inverted indexes of the content of the documents available in the legal information system (LIS). These systems can only handle syntactic aspects of the indexed documents, not the meaning.

One can assume that a syntactic connection between two documents indicates the existence of a semantic connection of some sort, i.e. the *meaning* of the content of the two documents may be related in some way. This observation can be used to propose an alternative way of searching for additional relevant documents: by following the paths that references to and from a particular document create to other documents in the system.

Documents and references create a structure, a *hyperstructure*. My task was looking at the possibilities to create an interface for searching that would make it possible to better utilize the syntactic connections between documents, e.g. by using an interface that would allow searching by navigation, instead of or in addition to the traditional searching by query formulation.

How to represent hyperstructures was an important question. My proposal was to use topic maps for the representation of hyperstructures. Topic Maps is an international standard (ISO/IEC 13250:2000) for representing “things” and “relations”.

Syntactic structures can to some degree be detected automatically by a computer. Identifying semantic structures is a process that requires human brain, it requires *understanding* the meaning of the text. To some extent it may be possible to detect *possible* semantic connections by performing a syntactic analysis of the texts.