Multilingual Ontological Analysis of European Directives

Gianmaria Ajani
Dipartimento di Scienze Giuridiche
Università di Torino - Italy
gianmaria.ajani@unito.it

Guido Boella
Leonardo Lesmo
Alessandro Mazzei
Dipartimento di Informatica
Università di Torino - Italy
[guido|lesmo|mazzei]@di.unito.it

Piercarlo Rossi
Dipartimento di Studi per l'Impresa e il Territorio
Università del Piemonte Orientale - Italy
piercarlo.rossi@eco.unipmn.it

Abstract

This paper describes the main features of our tool called "Legal Taxonomy Syllabus". The system is an ontology based tool designed to annotate and recover multi-lingua legal information and build conceptual dictionaries on European Directives.

1 Introduction

The European union each year produces a large number of Union Directives (EUD), which are translated into each of the communitary languages. The EUD are sets of norms that have to be implemented by the national legislations. The problem of multilinguism in European legislation has recently been addressed by using linguistic and ontological tools, e.g. (Boer et al., 2003; Giguet and P.S., 2006; Després and Szulman, 2006). The management of EUD is particularly complex since the implementation of a EUD however not correspond to the straight transposition into a national law. An EUD is subject to further interpretation, and this process can lead to unexpected results. Comparative Law has studied in details the problematics concerning EUD and their complexities. On the other hand managing with appropriate tools this kind of complexity can facilitate the comparison and harmonization of national legislation (Boer et al., 2003). Based on this research, in this paper, we describe the tool for building multilingual conceptual dictionaries we developed for representing an analysing the terminology and concepts used in EUD.

The main assumptions of our methodology, motivated by studies in comparative law (Rossi and Vogel, 2004) and ontologies engineering (Klein, 2001), are the following ones: 1) Terms and concepts must be distinguished; for this purpose, we use lightweight ontologies, i.e. simple taxonomic structures of primitive or composite terms together with associated definitions. They are hardly axiomatized as the intended meaning of the terms used by the community is more or less known in advance by all members, and the ontology can be limited to those structural relationships among terms that are considered as relevant (Oberle, 2005)[1]. 2) We distinguish the ontology implicitly defined by EUD, the *EU level*, from the various national ontologies, the *national level*. Furthermore, each national legislation refers to a distinct national legal ontology. We do not assume that the transposition of an EUD introduces automatically in a national ontology the same concepts present at the EU level. 3) Corresponding concepts at the EU level and at the national level can be denoted by different terms in the same national language.

In this paper, we show how the Legal Taxonomy Syllabus (LTS) is used to build a dictionary of consumer law, to support the Uniform Terminology Project[2] (Rossi and Vogel, 2004). The structure of this paper is the following one. In Section 2 we stress two main problems which comparative law has raised concerning EUD and their transpositions. In Section 3 we describe how the methodology of the LTS allows to cope with these problems and finally in Section 4we give some conclusions.

[1] See http://cos.ontoware.org/
[2] http://www.uniformterminology.unito.it

Proceedings of the ACL 2007 Demo and Poster Sessions, pages 21–24,
Prague, June 2007.

2 Terminological and conceptual misalignment

Comparative law has identified two key points in dealing with EUD, which makes more difficult dealing with the polysemy of legal terms: we call them the *terminological* and *conceptual misalignments*.

In the case of EUD (usually adopted for harmonising the laws of the Member States), the terminological matter is complicated by their necessity to be implemented by the national legislations. In order to have a precise transposition in a national law, a Directive may be subject to further interpretation. Thus, a same *legal concept* can be expressed in different ways in a Directive and in its implementing national law. The same legal concept in some language can be expressed in a different way in a EUD and in the national law implementing it. As a consequence we have a terminological misalignment. For example, the concept corresponding to the word *reasonably* in English, is translated into Italian as *ragionevolmente* in the EUD, and as *con ordinaria diligenza* into the transposition law.

In the EUD transposition laws a further problem arises from the different national *legal doctrines*. A legal concept expressed in an EUD may not be present in a national legal system. In this case we can talk about a conceptual misalignment. To make sense for the national lawyers' expectancies, the European legal terms have not only to be translated into a sound national terminology, but they need to be correctly detected when their meanings are to refer to EU legal concepts or when their meanings are similar to concepts which are known in the Member states. Consequently, the transposition of European law in the parochial legal framework of each Member state can lead to a set of distinct national legal doctrines, that are all different from the European one. In case of consumer contracts (like those concluded by the means of distance communication as in Directive 97/7/EC, Art. 4.2), the notion to provide in a *clear and comprehensible manner* some elements of the contract by the professionals to the consumers represents a specification of the information duties which are a pivotal principle of EU law. Despite the pairs of translation in the language versions of EU Directives (i.e., *klar und verständlich* in German - *clear and comprehensible* in English - *chiaro e comprensibile* in Italian), each legal term, when transposed in the national legal orders, is influenced by the conceptual filters of the lawyers' domestic legal thinking. So, *klar und verständlich* in the German system is considered by the German commentators referring to three different legal concepts: 1) the print or the writing of the information must be clear and legible (*gestaltung der information*), 2) the information must be intelligible by the consumer (*formulierung der information*), 3) the language of the information must be the national of consumer (*sprache der information*). In Italy, the judiciary tend to control more the formal features of the concepts 1 and 3, and less concept 2, while in England the main role has been played by the concept 2, though considered as plain style of language (not legal technical jargon) thanks to the historical influences of plain English movement in that country.

Note that this kind of problems identified in comparative law has a direct correspondence in the ontology theory. In particular Klein (Klein, 2001) has remarked that two particular forms of ontology mismatch are *terminological* and *conceptualization* ontological mismatch which straightforwardly correspond to our definitions of misalignments.

3 The methodology of the Legal Taxonomy Syllabus

A standard way to properly manage large multilingual lexical databases is to do a clear distinction among terms and their interlingual acceptions (or *axies*) (Sérasset, 1994; Lyding et al., 2006). In our system to properly manage terminological and conceptual misalignment we distinguish in the LTS project the notion of legal term from the notion of legal concept and we build a systematic classification based on this distinction. The basic idea in our system is that the conceptual backbone consists in a taxonomy of concepts (ontology) to which the terms can refer to express their meaning. One of the main points to keep in mind is that we do not assume the existence of a single taxonomy covering all languages. In fact, it has been convincingly argued that the different national systems may organize the concepts in different ways. For instance, the term *contract* corresponds to different concepts

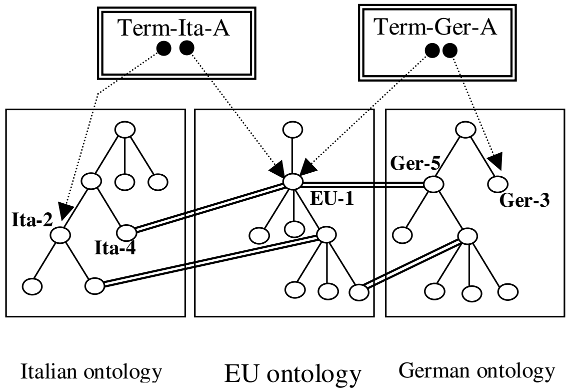

Figure 1: Relationship between ontologies and terms. The thick arcs represent the inter-ontology "association" link.

in common law and civil law, where it has the meaning of *bargain* and *agreement*, respectively (Sacco, 1999). In most complex instances, there are no homologous between terms-concepts such as *frutto civile* (legal fruit) and *income*, but respectively civil law and common law systems can achieve functionally same operational rules thanks to the functioning of the entire taxonomy of national legal concepts (Graziadei, 2004). Consequently, the LTS includes different ontologies, one for each involved national language plus one for the language of EU documents. Each language-specific ontology is related via a set of *association* links to the EU concepts, as shown in Fig. 1.

Although this picture is conform to intuition, in LTS it had to be enhanced in two directions. First, it must be observed that the various national ontologies have a reference language. This is not the case for the EU ontology. For instance, a given term in English could refer either to a concept in the UK ontology or to a concept in the EU ontology. In the first case, the term is used for referring to a concept in the national UK legal system, whilst in the second one, it is used to refer to a concept used in the European directives. This is one of the main advantages of LTS. For example *klar und verständlich* could refer both to concept `Ger-379` (a concept in the German Ontology) and to concept `EU-882` (a concept in the European ontology). This is the LTS solution for facing the possibility of a correspondence only partial between the meaning of a term has in the national system and the meaning of the same term in the translation of a EU directive. This feature enables the LTS to be more precise about what "translation" means. It puts at disposal a way for asserting that two terms are the translation of each other, but just in case those terms have been used in the translation of an EU directive: within LTS, we can talk about direct EU-translations of terms, but only about indirect national-system translations of terms. The situation enforced in LTS is depicted in Fig. 1, where it is represented that: The Italian term *Term-Ita-A* and the German term *Term-Ger-A* have been used as corresponding terms in the translation of an EU directive, as shown by the fact that both of them refer to the same EU-concept `EU-1`. In the Italian legal system, *Term-Ita-A* has the meaning `Ita-2`. In the German legal system, *Term-Ger-A* has the meaning `Ger-3`. The EU translations of the directive is correct insofar no terms exist in Italian and German that characterize precisely the concept `EU-1` in the two languages (i.e., the "associated" concepts `Ita-4` and `Ger-5` have no corresponding legal terms). A practical example of such a situation is reported in Fig. 2, where we can see that the ontologies include different types of arcs. Beyond the usual *is-a* (linking a category to its supercategory), there are also a *purpose* arc, which relates a concept to the legal principle motivating it, and *concerns*, which refers to a general relatedness. The dotted arcs represent the reference from terms to concepts. Some terms have links both to a National ontology and to the EU Ontology (in particular, *withdrawal* vs. *recesso* and *difesa del consumatore* vs. *consumer protection*).

The last item above is especially relevant: note that this configuration of arcs specifies that: 1) *withdrawal* and *recesso* have been used as equivalent terms (concept `EU-2`) in some European Directives (e.g., Directive 90/314/EEC). 2) In that context, the term involved an act having as purpose the some kind of protection of the consumer. 3) The terms used for referring to the latter are *consumer protection* in English and *difesa del consumatore* in Italian. 4) In the British legal system, however, not all *withdrawals* have this goal, but only a subtype of them, to which the code refers to as *cancellation* (concept `Eng-3`). 5) In the Italian legal system, the term *diritto di recesso* is ambiguous, since it can be used with reference either to something concerning

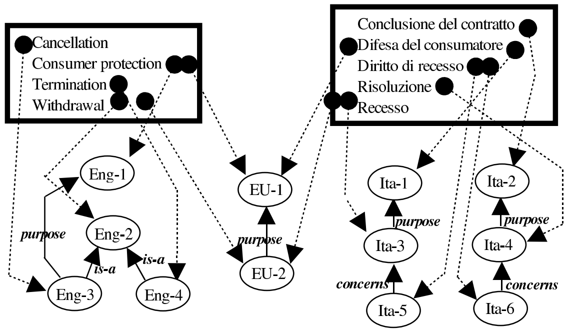

Figure 2: An example of interconnections among terms.

the *risoluzione* (concept `Ita-4`), or to something concerning the *recesso* proper (concept `Ita-3`).

Finally, it is possible to use the LTS to translate terms into different national systems via the concepts which they are transposition of at the European level. For instance suppose that we want to translate the legal term *credito al consumo* from Italian to German. In the LTS *credito al consumo* is associated to the national umeaning `Ita-175`. We find that `Ita-175` is the transposition of the European umeaning `EU-26` (*contratto di credito*). `EU-26` is associated to the German legal term *Kreditvertrag* at European level. Again, we find that the national German transposition of `EU-26` corresponds to the national umeaning `Ger-32` that is associated with the national legal term *Darlehensvertrag*. Then, by using the European ontology, we can translate the Italian legal term *credito al consumo* into the German legal term *Darlehensvertrag*.

4 Conclusions

In this paper we discuss some features of the LTS, a tool for building multilingual conceptual dictionaries for the EU law. The tool is based on lightweight ontologies to allow a distinction of concepts from terms. Distinct ontologies are built at the EU level and for each national language, to deal with polysemy and terminological and conceptual misalignment.

Many attempts have been done to use ontology in legal field, e.g. (Casanovas et al., 2005; Després and Szulman, 2006) and LOIS project (that is based on EuroWordNet project (Vossen et al., 1999), `http://www.loisproject.org`), but to our knowledge the LTS is the first attempt which starts from fine grained legal expertise on the EUD domain.

Future work is to study how the LTS can be used as a thesaurus for general EUD, even if the current domain is limited to consumer law.

References

A. Boer, T.M. van Engers, and R. Winkels. 2003. Using ontologies for comparing and harmonizing legislation. In *ICAIL*, pages 60–69.

P. Casanovas, N. Casellas, C. Tempich, D. Vrandecic, and R. Benjamins. 2005. OPJK modeling methodology. In *Proceedings of the ICAIL Workshop: LOAIT 2005*.

S. Després and S. Szulman. 2006. Merging of legal micro-ontologies from european directives. *Artificial Intelligence and Law*, ??:??–?? In press.

E. Giguet and P.S. 2006. Multilingual lexical database generation from parallel texts in 20 european languages with endogenous resources. In *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions*, pages 271–278, July.

M. Graziadei. 2004. Tuttifrutti. In P. Birks and A. Pretto, editors, *Themes in Comparative Law*, pages –. Oxford University Press.

M. Klein. 2001. Combining and relating ontologies: an analysis of problems and solutions. In *Workshop on Ontologies and Information Sharing, IJCAI'01*, Seattle, USA.

V. Lyding, Elena Chiocchetti, G. Sérasset, and F. Brunet-Manquat. 2006. The LexALP information system: Term bank and corpus for multilingual legal terminology consolidated. In *Proc. of the Wokshop on Multilingual Language Resources and Interoperability, ACL06*, pages 25–31.

D. Oberle, editor. 2005. *Semantic Management of Middleware*. Springer Science+Business and Media.

P. Rossi and C. Vogel. 2004. Terms and concepts; towards a syllabus for european private law. *European Review of Private Law (ERPL)*, 12(2):293–300.

R. Sacco. 1999. Contract. *European Review of Private Law*, 2:237–240.

G. Sérasset. 1994. Interlingual lexical organization for multilingual lexical databases in NADIA. In *Proc. COLING94*, pages 278–282.

P. Vossen, W. Peters, and J. Gonzalo. 1999. Towards a universal index of meaning. In *Proc. ACL-99 Siglex Workshop*.