

Multilingual Conceptual Dictionaries Based on Ontologies

Gianmaria Ajani*, Leonardo Lesmo[°], Guido Boella[°], Alessandro Mazzei[°], Piercarlo Rossi*

* *Dipartimento di Scienze Giuridiche, Università di Torino - Italy -
gianmaria.ajani@unito.it. p.rouge@inwind.it*

[°] *Dipartimento di Informatica, Università di Torino - Italy -
{lesmo,guido,mazzei}@di.unito.it*

Abstract. This paper introduces a new tool called “Legal Taxonomy Syllabus”. This is an ontology based tool designed in order to annotate and recover multi-lingua legal information.

Keywords: Multilingual Legal Ontologies

1. Introduction

The legal orders of the EC Member States are currently on the verge of further convergence of their respective private law - a process which is strongly influenced and directed by European primary and secondary legislation.

Several researches in the field of European law are currently targeting the identification of similarities and differences, common principles and concepts within national and European legal institutions. A EU legal terminology is however required to properly represent the European concepts and differentiate them from those of Member States [European Commission, 2003a, 2004].

1. That is a priority identified by the European Commission: “rewriting legal texts” is necessary “to render them more coherent and understandable” [European Commission, 2003b], in order to enhance a uniform interpretation of EU law in each national legal system. Terminological fragmentation causes lack of internal and external coherence in European law, especially where the impact of EU legislation is significant such as in a subject matter like consumer protection [European Commission, 2001]. The internal coherence is hindered by the potentially legal uncertainty resulting from inconsistent definitions of EU legal terms and from the contradictory use of legal terms within different sectoral legislative interventions (see EC Directives on consumer law, like Timeshare, Distant Contract, Unfair terms, and so forth). The external coherence is frustrated by the differences in legal conceptualisations that may import the “result of Member State transposition which has itself added unnecessary, complicated, detailed or excessive

provisions” [European Commission, 2003b]. This is the case of Directives that need to be implemented by national statutory instruments. The implementation of a Directive may not correspond to its straight transposition in a national law and may be subject to further interpretation. Thus a same legal concept can be expressed in different ways in a Directive and in the transposition law. For example, in Directive 99/44/EC, the concept corresponding to the English word “reasonably” is translated into Italian with “ragionevolmente” in the language of the Directive, and with “con ordinaria diligenza” in the Italian transposition law. We can define this problematic issue as a sort of extra-EU polysemy. This implementation process is problematical, since it leads to use a different terminology in the EU legislation and in the national legislation.

Moreover, the terminological fragmentation is considered by the European Commission as an obstacle to “developing more user friendly access to consult and use Community law” [European Commission, 2003b]. Consider for example EUR-Lex, a wide portal providing access to all the official legal documents of the European Union, though with the paradox that, in order to obtain a full coverage, it limits the complete accessibility to legal documents, particularly for the needs of lawyers.

Reporting too large instances for each query without comprehensible classifications for the expectances of national jurists and practitioners hinders the applicability of EUR-Lex for most legal uses in the Member States’ legal orders. Where the searched legal terms do not properly correspond to the given legal terms in the EUR-Lex database, the queries submitted by the lawyers become enlarged to ordinary language or full text access with several limitations in the possibility of managing such an amount of data.

2. The criteria for a comparison of legal concepts among different legal systems deal with the question of what legal concepts are and of which is their relation with the terms representing them. It is clear that the problem of the legal concepts which sometimes emerges from the terminological transposition of the Directives is wider and constitutive of the European legal order, where several national legal systems with their own traditions have been converging in a new European order that does not substitute them but interplays with them.

This state of affairs represents an opportunity to demonstrate how information technology can prevent the problem of knowledge representation known as conceptual misalignment.

In fact, the difference in terminology sometimes corresponds to a difference in meaning and sometimes does not. Moreover, some legal traditions prefer to use different terms for different concepts while oth-

ers do the opposite. Unfortunately, that problem is underestimated in the process of drafting European law and implementing it in national instruments as well as in the classification systems of main European legal database, such as EUR-Lex. So, the German phrase “Klar und verständlich”, used in several Directives on consumer law, is ambiguous between the concepts denoted in Italian respectively by “chiaramente”, “chiaro e comprensibile” and “inequivoco”, but in the Italian language version of the same Directive that difference is not traced back.

One tool which might help to increase European terminological consistency taking into account the problem of knowledge representation is the Legal Taxonomy Syllabus which has already set off two years ago by the Department of Computer Science of the University of Turin specialised in ontology for the web domain, legal ontology and AI studies on law, which cooperates with the main groups of research on this topic in the international context, such as the LOA, Laboratory for Applied Ontology of Italian CNR, CIRSIFID at the University of Bologna, and so forth.

Such tool has been developed to support the Uniform Terminology project [Rossi and Vogel, 2004], in the overall collaboration with the Law Department of Turin, specialised in comparative and European law and involved in several substantial researches funded by the Commission on: a) legal terminology, like the Uniform Terminology for European Private Law - EU Fifth Framework Program - “Improving Human Potential” (<http://www.uniformterminology.unito.it>); b) common frame of reference, like the Joint Network on European Private Law - EU Sixth Framework Program “Network of Excellence” (<http://www.copecl.org>).

Legal Taxonomy Syllabus is designed as an open-access database linking European terms with national transposition law and also linking terms horizontally (i.e. between national legal orders). It provides full text reference of relevant EU and Member States’ legislation. The database includes related case law and short commentary notes by national scholars where this is necessary to describe differing legal doctrine. As a starting point, the Legal Taxonomy Syllabus covers consumer law with national law references limited to France, Germany, Italy, Spain and the UK. The Legal Taxonomy Syllabus could be useful for lawyers, translators, legislators and scholars: the tool may help legislators to enhance terminological coherency already at drafting stage of legal acts. The cross-reference features enable lawyers to search for relevant case law in other Member States by a one-click method. Unlike a dictionary, the Legal Taxonomy Syllabus does not only provide translators with a proper translation but provides the respective legislative context of each term.

The database can also help scholars to perform cross-sector analyses on the use of legal terms and concepts, integrating with the ontologies the classification systems employed by existing database, which are oriented towards the multilingualism terminology but only in one dimension, that of the European Union; or which are structured on the subdivisions and sections of the European Treaty that have no relevant connection with the classifications implied in the legal discourse within the Member States.

What seems important to us is to introduce the Legal Taxonomy Syllabus system among the developed tools, as a new way to properly manage the diverseness of EU law. The treatment of legal terminology adopts a mixed descriptive-prescriptive terminological application to the corpora constituted by the European legal documents. In this way, prior to improving the consistency of the EU terminology, the research is focused on the understanding of such a terminology within the Member States, highlighting the matter of polysemy and amphibology of the terms used in a same language (such as French) at the European level and at the national level of France or Belgium.

2. From terms to ontologies: the Legal Taxonomy Syllabus system

The tool that we propose is based on a clear distinction between the notions of “legal term” and “legal concept”. The basic idea is that the basic 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 in common law and civil law, where it has the meaning of “bargain” and “agreement”, respectively argued [Sacco, 1999, Pozzo, 2003].

Consequently, the Legal Taxonomy Syllabus includes different ontologies, one for each involved 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.

2.1. POLYSEMY AND TRANSLATIONS

Although this picture is conform to intuition, in the Legal Taxonomy Syllabus it had to be enhanced in two directions. First, it must be

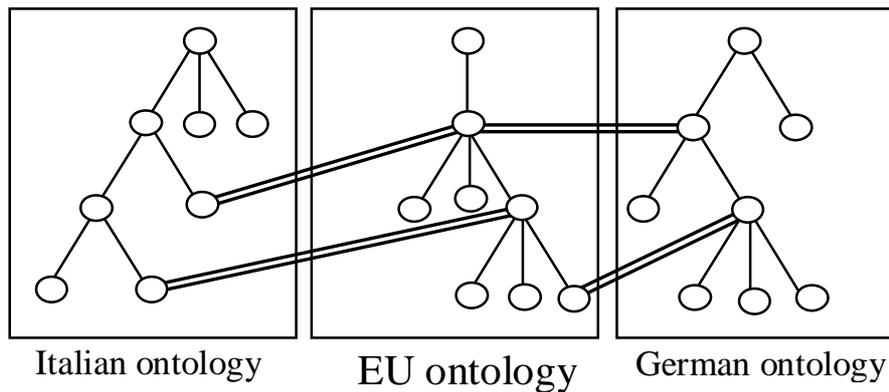


Figure 1. Relationship between ontologies. The thick arcs represent the inter-ontology “association” link.

observed that the various national ontologies have a reference language. This is not the case for the EU ontology. In fact, a given term in, say, 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 the Legal Taxonomy Syllabus. For example “Klar und verständlich” could refer both to concept “De379” (a concept in the German Ontology) and to concept “EU882” (a concept in the European ontology). This is the Legal Taxonomy Syllabus solution for facing the possibility of a correspondence only partial between the meaning 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 Legal Taxonomy Syllabus 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 the Legal Taxonomy Syllabus, we can talk about direct EU-translations of terms, but only about indirect national-system translations of terms. The situation enforced in the Legal Taxonomy Syllabus is depicted in fig.2, 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

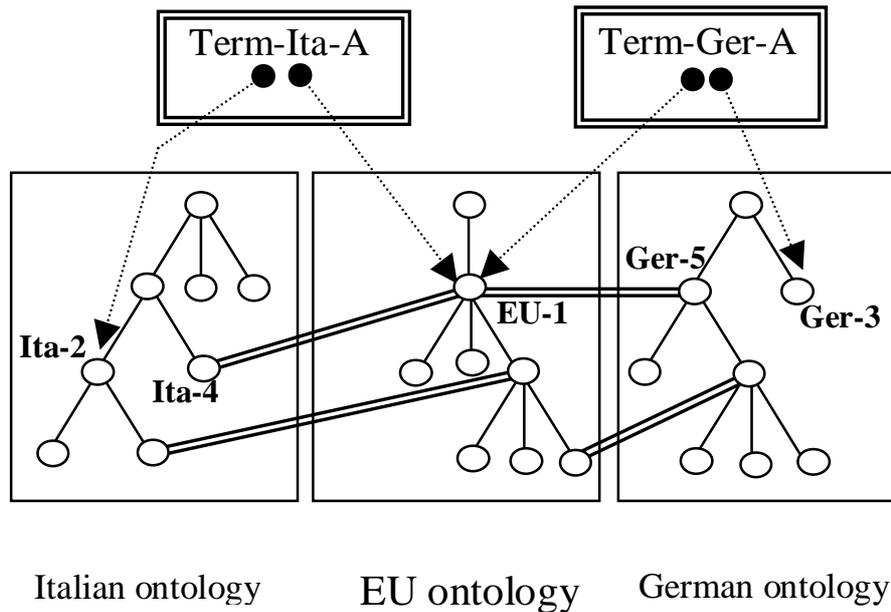


Figure 2. Relationship between ontologies and terms

precisely the concept EU-1 in the two languages (i.e the “associated” concepts Ita-4 and Ita-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 type of arcs. Beyond the standard is-a (linking a category to its supercategory), there is also a purpose arc, which is self-explanatory The dotted arc 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. “Diritto di 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 “Diritto di Recesso” have been used as equivalent terms (concept EU-2) in some European Directive. 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” also refers to a kind of “risoluzione” (concept Ita-3).

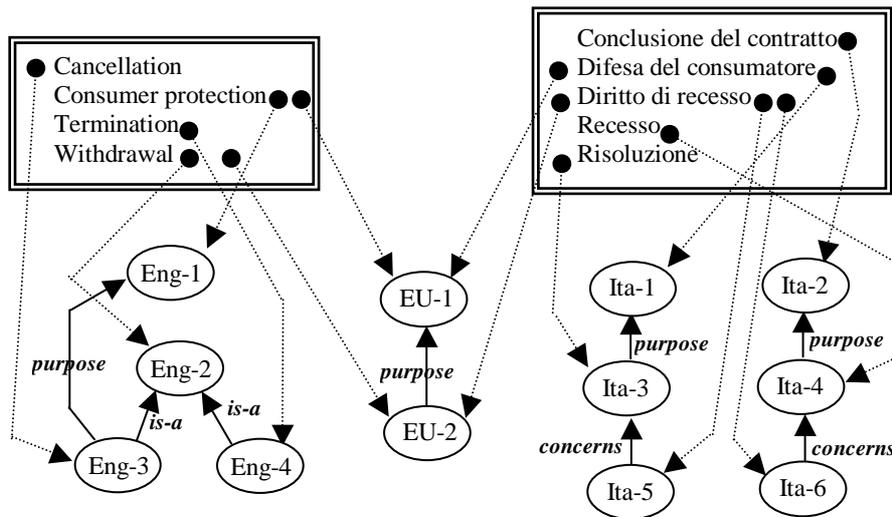


Figure 3. An example of interconnections among terms

All of this seem to correspond neatly to the conception of terminology that is currently accepted by the scholars in comparative law. For instance, it can safely be stated that the term “diritto di recesso” as used in consumer law directives (i.e. the right of a consumer to withdraw from a contract) does not correspond to an existing legal concept in the Italian legislation. Moreover, the right of withdrawal appearing in EU directives also differs from the doctrine with the same label of the British system.

2.2. THE LEGAL TAXONOMY SYLLABUS KNOWLEDGE BASE

This complex scenario shows how the traditional top-down approach to the development of legal ontologies [Visser and Bench-Capon, 1998] is not flexible enough. Usually, ontologies are built starting from very general concepts which are then specialised in more detailed concepts. Moreover most ontologies are oriented to a single national tradition. In this process the knowledge engineers risk not to take into account the interpretation process of the legal specialists on the real multilingual data. These ontologies aim at modelling the legal code but not the legal “doctrine”, that is the work of interpretation and re-elaboration of the legal code which is fundamental for transposing EUD into national laws.

In the development of the ontologies described in the previous section, we propose to follow the approach of the “Uniform terminology project” Ajani and Ebers [2005]. As a first step, terms are collected in a

database together with the legal sources where they appear, in order to identify the concepts. Then, for each EU language, the set of concepts is organized in an ontology which can be different for different legal traditions. This reconstruction work is done by legal experts rather than knowledge engineers. In this phase the result is a light-weight ontology rather than an axiomatic one. Only relations among terms are identified without introducing restrictions and axioms. The function of these ontologies is to compare the taxonomic structure in the different legislations, to provide a form of intelligent indexing and to draw new legal conclusions.

In a second phase, a knowledge engineer can reorganize the ontology and integrate it with a top-level well-founded ontology like DOLCE [Gangemi et al., 2002].

3. Accessing to the Legal Taxonomy Syllabus

Even if many tools for the construction of ontologies are available (e.g., Protegè), we had to design a new development system based on the constraints of being *distributed* and *user-friendly*. As described in Section 2, the central step of the legal concepts annotation is performed by legal experts from different countries rather than knowledge engineer. As a consequence, the Legal Taxonomy Syllabus has been designed by using a very simple client-server web application¹. The ontology framework has been inspired by the Gene Ontology project (<http://www.geneontology.org/>), from which it inherits the logical and graphical representation. The tool has two levels of use.

In the first level the web interface makes available to the legal expert a friendly way to introduce data about terms and concepts. At this level, the tool aims at providing the expert with a support in her/his activity of term comparison, Fig. 4. In other words, instead of using a standard database interface, the expert can specify the correspondences among terms found via the manual inspection of EU directives or ECJ decisions (or in national legislation and case-law) in a controlled way and save in an ontology structure the result of her/his analysis.

The second level is devoted to a user who wants to retrieve the documents related to a given legal term. At this level, Legal Taxonomy Syllabus acts as the desired extension of a standard legal database (cf. “EUR-Lex”, see below), by enabling the user to find the relevant documents taking into account the complex net of semantic correspon-

¹ The web server has been implemented by using a WAPP (Windows, Apache, PostgreSQL, PHP) platform.



Figure 4. A screenshot of the Legal Taxonomy Syllabus interface for navigation and update of the ontologies.

dences that characterize the relationships between legal terms at the international level (Fig. 5).

The importance of such two levels of use has ever not been stressed sufficiently. As pointed out in jurisprudence and European law literature [Patterson, 1996, Gerven, 2000], the classification schemes of legal objects (such as the distinction between rights *in rem* and rights *in personam* with the reference to the time-share property) does not exist externally to the legal domain and exists only because the legal science deems it does. The principle of consensus (the intentional use of the majority of scholars operating within the relevant discipline) and not only the interoperability may foster the ontological knowledge at the level of legal domain. Consequently, if many classification schemes are adopted in several legal orders the ontology should take all into account before refining them.

3.1. TOWARDS XML FORMAT

From technological point of view, the Legal Taxonomy Syllabus knowledge is stored in a relational database, i.e. in a number of SQL tables. As

The screenshot shows the 'Legal Taxonomy Syllabus' interface. The browser address bar displays 'http://normas.di.unito.it/syllabus/index.php?action=loadUmeaning&umeaningid=84'. The page title is 'Legal Taxonomy Syllabus' with the subtitle 'A Dictionary and Repository on European Law'. The interface includes a navigation menu with options like 'Home', 'Database', 'Commentary', 'Howto', 'Members', 'Intra', and 'Links'. A sidebar on the left contains various administrative and user-related links such as 'Logout', 'Term search', 'Umeaning Management', 'Relations', 'Legal languages', 'Users', 'Log', 'Database Maintenance', and 'User's manual (italian)'. The main content area is titled 'Ontology' and shows an 'Ontology Graph' with a relation 'hypon' between 'time of time' and 'time of conclusion of the contract'. Below this, a table displays the 'European level' details for the term 'time of conclusion of the contract'.

Actions	Legal language	Term	References	Notes	Attachments	Associations
Edit Delete Insert synonym Insert another linguistic version Insert attachment Insert transposition Add child Add hypernym Add hyponym Add quasi-synonymous	English	time of conclusion of the contract	Directive 1985/577/EEC, Article 4 In the case of transactions within the scope of Article 1, traders shall be required to give consumers written notice of their right of cancellation within the period laid down in Article 5, together with the name and address of a person against whom that ...	ECJ Case Law: [...] Notes: [...]		time of conclusion of the contract (delete transposition)

Umeaning created by Christian Vogel at 01 June 2006 alle ore 10:13:57
Last modification: 16 June 2006 alle ore 12:33:08

Figure 5. A screenshot of the Legal Taxonomy Syllabus interface for searching a legal term.

we have pointed out in the previous section, the database structure used in the Legal Taxonomy Syllabus is a descendant of the structure used by the Gene Ontology database. Gene ontology project includes a number of programs that allows us the transformation from the SQL tables into a number of formats. In particular these program allow the conversion into the OWL format. Anyway there are a number of differences between the Legal Taxonomy Syllabus and the Gene ontology tables, e.g. the latter do not take the concept-term distinction into account. In the next future work we intend to modify the programs provided by Gene ontology project in order to work fine with the Legal Taxonomy Syllabus SQL relational schema. The modification has to account for three critical issues deriving from the term-concept distinction: 1. one single term can correspond to a number of distinct concepts. 2. one single concept can correspond to a number of distinct terms belonging to the same language. 3. one single concept can correspond to a number of distinct terms belonging to different languages.

4. Conclusion and future development

There is a number of works that consider the theoretical issues related to the construction of legal ontologies McCarty [1989], Stamper [1991], Breuker et al. [1997]. In particular the framework presented in Kraligen [1997] is a frame-based system that classifies the legal facts. A basic component of this system is the *legal concept description*, i.e. Kraligen proposes a distinction between a legal term and a legal concept similar to the distinction that we have adopted in Legal Taxonomy Syllabus. From a practical point of view, there are two projects that are related in some way to the Legal Taxonomy Syllabus. The “EURLex” system (<http://europa.eu.int/eur-lex/>) is a web portal that interfaces a number of databases in order to access a wide collection of legal documents produced by the EU. However, in order to obtain a full coverage, EURLex limits the complete accessibility to legal documents, particularly for the needs of lawyers. Each query, even when using boolean search, reports too large instances without comprehensible classifications for the expectancies of national jurists and practitioners, and thus hinders the applicability of EUR-Lex for most legal uses in the Member States’ legal.

“Eurovoc” (<http://europa.eu.int/celex/eurovoc/>) is a web application that accesses a number of multilingual thesauri. The main point of this project is the splitting of the legal terms into two sets: the *descriptor* and *non-descriptor*. A non-descriptor legal term can always be mapped into a descriptor legal term that has the same meaning. Moreover, the basic hypothesis is that each descriptor can be translated straightforwardly into the official languages of the EU. In contrast to the Legal Taxonomy Syllabus, the main purpose of Eurovoc is the information extraction. Indeed, the sparseness problems related to the *bags of word* techniques can be reduced by replacing the non-descriptor with the corresponding descriptor. However Eurovoc does not distinguish between a legal term and a legal concept, and cannot resolve easily the problems related to the polysemy.

“LOIS” Project (<http://www.loisproject.org>) aims at extending EuroWordnet with legal information. Whilst the final goal of LOIS is to support applications concerning information extraction, the Legal Taxonomy Syllabus we propose herein is concerned with the access of human experts to the EU documents.

In this paper we have described the Legal Taxonomy Syllabus. The actual ongoing phase of this project involves the collection of the legal terms and legal concepts by a group of legal experts belonging

to a number different countries (in particular Italy, England, Germany, Spain, France). In the next work we intend to release new tools to allow the conversion from the SQL tables, used by the Legal Taxonomy Syllabus database, into the OWL format.

Acknowledgements

The Legal Taxonomy Syllabus project has been partially funded by the Uniform Terminology project (www.uniformterminology.unito.it, Contract n. HPRN-CT-2002-00229).

References

- G. Ajani and M. Ebers, editors. *Uniform Terminology for European Contract Law*. Nomos, Baden Baden, 2005.
- J. Breuker, A. Valente, and R. Winkels. Legal ontologies: A functional view. In *Procs. of 1st LegOut Workshop on Legal Ontologies*, pages 23–36, 1997.
- European Commission. Green Paper on EU Consumer Protection. *Communication from the Commission to the European Parliament and the Council*, COM(2001)(531), 2001.
- European Commission. More Coherent European Contract Law, An Action Plan. *Communication from the Commission to the European Parliament and the Council*, COM(2003)(68 final), 2003a.
- European Commission. Updating and simplifying the Community acquis. *Communication from the Commission*, COM(2003)(71 final), 2003b.
- European Commission. European Contract Law and the revision of the acquis: the way forward. *Communication from the Commission*, COM(2004)(651 final), 2004.
- A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening ontologies with dolce. In *Proceedings of EKAW*, pages 21–29, 2002.
- W. Van Gerven. Of rights, remedies and procedures. *Common Market Law Review*, pages 501–536, 2000.

- V. Kraligen. A conceptual frame-based ontology for the law. In *Procs. of 1st LegOut Workshop on Legal Ontologies*, pages 15–22, 1997.
- L.T. McCarty. A language for legal discourse: Basic features. In *Proc. of Second International Conference on Artificial Intelligence and Law*, 1989.
- D. Patterson. *Law and Truth*. Oxford University Press, 1996.
- B. Pozzo. Harmonisation of european contract law and the need of creating a common terminology. *European Review of Private Law*, 6:754–767, 2003.
- P. Rossi and C. Vogel. Terms and concepts; towards a syllabus for european private law. *European Review of Private Law (ERPL)*, 12 (2):293–300, 2004.
- R. Sacco. Contract. *European Review of Private Law*, 2:237–240, 1999.
- R.K. Stamper. The role of semantics in legal expert systems and legal reasoning. *Ratio Juris*, 4(2):219–244, 1991.
- P.R.S. Visser and T.J.M. Bench-Capon. A comparison of four ontologies for the design of legal knowledge systems. *Artificial Intelligence and Law*, 6:27–57, 1998.

