



## UvA-DARE (Digital Academic Repository)

### Ontology Representation : design patterns and ontologies that make sense

Hoekstra, R.J.

**Publication date**  
2009

[Link to publication](#)

#### **Citation for published version (APA):**

Hoekstra, R. J. (2009). *Ontology Representation : design patterns and ontologies that make sense*. [Thesis, fully internal, Universiteit van Amsterdam]. IOS Press.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

---

## Chapter 6

# Commonsense Ontology

### 6.1 Introduction

In many ways, the corpus of legal information available today is the world wide web's little sister, at least qua structure (Hoekstra, 2008).<sup>1</sup> It consists of a huge volume of heterogeneous, but closely inter-linked documents. These documents are increasingly being made available in digital form as part of public accessibility projects by governments.<sup>2</sup> However, a major difference is that the relations between legal texts are typically expressed in natural language only. Also, these references are not always absolute, typically point to *parts* of documents, and often import an externally defined meaning of a term (de Maat et al., 2008; de Maat et al., 2006). Consolidation of such semantic references into a single representation introduces a significant maintenance issue, as legal texts are very dynamic and undergo change independently from each other. In fact, the meaning of terms in law imposes an ordering on entities in reality that can change over time, but *stays applicable* to older cases. In short, law adopts an intricate versioning scheme (Boer et al., 2004a,b; Klarman et al., 2008). The MetaLex/CEN<sup>3</sup> XML standard for legislative sources provides an XML schema for representing the structure and dynamics of legal texts (Boer et al., 2007c, 2002).

A semantic representation – be it for the purpose lightweight annotation, consistency checking or legal knowledge based reasoning (planning, assessment) – should take the dynamic and structural properties of legal texts into account. This is most directly reflected in the principle of *traceability*: any representation of some legal text should be traceable to its original source; it should be *isomorphic* (Bench-Capon and Coenen, 1991). A representation of some (part of) legislation is dependent on that legislation, and is therefore essentially al-

---

<sup>1</sup>This chapter is a revised version of a series of papers on the LKIF Core ontology written together with Joost Breuker and Alexander Boer (Breuker et al., 2006, 2007; Hoekstra et al., 2007, 2008).

<sup>2</sup>An example is the portal of the Dutch government <http://www.wetten.nl> that discloses currently active legislation

<sup>3</sup>CEN is the European Committee for Standardisation, See <http://www.metalex.eu>, <http://legacy.metalex.eu> and <http://www.cen.eu>

ways an *annotation* on that text. This view is central to e.g. the *structure blocks* of van Engers and Glassée (2001), where the UML/OCL representation of structural elements in legislation is organised in corresponding UML packages.<sup>4</sup>

The MetaLex/CEN initiative provides a standard transformation of XML encoded legal texts to RDF/XML. More elaborate, formal representations of the contents of the texts in OWL are then related to this RDF representation. A relatively uncharted application area for this approach is that of *compliance*, where the business processes of organisations (businesses and governments alike) need to be aligned with respect to some body of regulations. An additional requirement of the legal domain is that definitions of concepts should be represented in such a way that their semantic interpretation mimics the structure and applicability of the texts. This includes means to *scope* definitions with respect to particular *parts of a text*, as in e.g. deeming provisions, regarding the *temporal validity* of a text (Klarman et al., 2008), and concerning its *jurisdiction* (Boer et al., 2005b, 2007b).

One could argue that such requirements indicate the need for knowledge representation languages specific to law (as in e.g. deontic logics). However, the legal field is in one respect wholly analogous to the web in that legal information is used and incorporated in a wide variety of systems, each using the information in different ways. Also, the whole body of legal information is not maintained by a single issuer, but rather by a significant number of authorities that each publish, incorporate, extend, comply with, enforce and implement regulations. Therefore, the requirements for knowledge representation on the Semantic Web hold for representation of legal sources as well. Especially as the information exchange between those parties can benefit enormously from a well designed standard. This principle lies at the heart of the Legal Knowledge Interchange Format (LKIF) that allows for the interchange of legal knowledge between commercial vendors (Boer et al., 2007a,c).<sup>5</sup>

As an interchange format, one would expect LKIF to be a knowledge representation language in its own right just as e.g. the KIF of Genesereth and Fikes (1992) or its successor CommonLogic (ISO/IEC, 2007). However, LKIF rather specifies a knowledge interchange *architecture* for alternative ways of expressing different types of legal knowledge using three different representation paradigms. The first approach is characterised by the *pragmatism* of legal knowledge based system vendors. Commercial representation languages are typically rather inexpressive rule formalisms and constructed for the specific purpose of supporting features of the vendor's application suite (Gordon et al., 2007b). The second paradigm is defined by a theoretical focus on the *epistemological status* of knowledge as part of legal reasoning. In this view, a legal knowledge representation language does not purport to provide definitions, but rather characterises the use and manipulation of information to establish and *justify* legal facts. Legal knowledge is interpreted as the knowledge of reaching a legally valid conclusion. In LKIF, this strategic *control* knowledge (Clancey, 1983; Breuker and Van De Velde, 1994; van Heijst et al., 1997) is captured by a defeasible rule formalism, and more prominently by a theory of

<sup>4</sup>OCL: the Object Constraint Language, see <http://www.omg.org/technology/documents/formal/ocl.htm>.

<sup>5</sup>LKIF is developed as part of the Estrella project: European project for Standardized Transparent Representations in order to Extend Legal Accessibility (IST-2004-027655), see <http://www.estrellaproject.org>

argumentation (Boer et al., 2007a; Gordon et al., 2007a; Gordon, 2007).<sup>6</sup>

Where the second perspective emphasises epistemological status and legal theory – i.e. it does not include an explicit domain theory – the third approach adopts the *knowledge representation* perspective outlined in the preceding chapters. In this view, legal knowledge is characterised as that which is reasoned *with* rather than a specification of legal reasoning itself. The former can be expressed as a domain theory using the OWL DL knowledge representation formalism, while the latter is captured by meta components that implement problem solving methods over a monotonic DL knowledge base (Breuker and Van De Velde, 1994), as in the knowledge system *shells* of Marcus (1988) or Chandrasekaran and Johnson (1993). Typical legal reasoning tasks are *legal assessment*, i.e. determining whether some case or situation is allowed or disallowed given a system of normative statements, *legal planning* (Valente, 1995), but also *legal argumentation*. However, the way in which these tasks are performed is not necessarily particular to the legal domain. For instance, though legal argumentation differs from argumentation in general in that it is *adversarial* and takes place in court, the positing of arguments (hypotheses) and support (proofs, findings) is central to any type of problem solving (cf. Breuker (1994); Boer (2000), and Figure 2.9). The knowledge representation perspective allows for an isomorphic representation of the *contents* of legal sources as annotation, required for traceability and maintainability.

### 6.1.1 A Functional View

Clancey (1983) identified different types of rules involved in the medical domain theory of the MYCIN system (Section 2.3.2): identification, causal, world fact and domain fact rules. Similarly, the domain theory of *legal* knowledge based systems is not homogeneous either. Valente (1995); Valente and Breuker (1995) give a general breakdown of the types of knowledge and their dependencies involved in the legal domain. Valente’s Functional Ontology of Law (FOLaw) describes the legal system as an instrument to influence society and reach certain social goals, i.e. it exists to fulfil a *function*. The legal system can be viewed as a “social device operating within society and on society, and whose main function is to regulate social behaviour” (Valente, 1995, p. 49).

Despite its name, FOLaw is not really an ontology at all, at least not in the strict sense. It describes the categories of legal knowledge we can reason about, i.e. the kinds of things one can ‘know’ in the legal domain and *not* the things that ‘exist’. It is an *epistemological framework*, an ontology of the epistemology of law, cf. (Breuker and Hoekstra, 2004c, and Section 5.5.2). FOLaw distinguishes seven types of knowledge: commonsense knowledge, world knowledge, normative knowledge, responsibility knowledge, meta-legal knowledge, creative knowledge and reactive knowledge. Any legal knowledge based system will incorporate at least one of these knowledge types. In fact, FOLaw does not just have a functional perspective on the legal system itself, but on the different knowledge components within this system as well (Levesque, 1984, and Section 2.4.1).

The most characteristic category of legal knowledge is *normative knowledge*.

<sup>6</sup>A reasoner for unfolding argumentation schemes is implemented as part of the Carneades argument mapping system, <http://carneades.berlios.de/>.

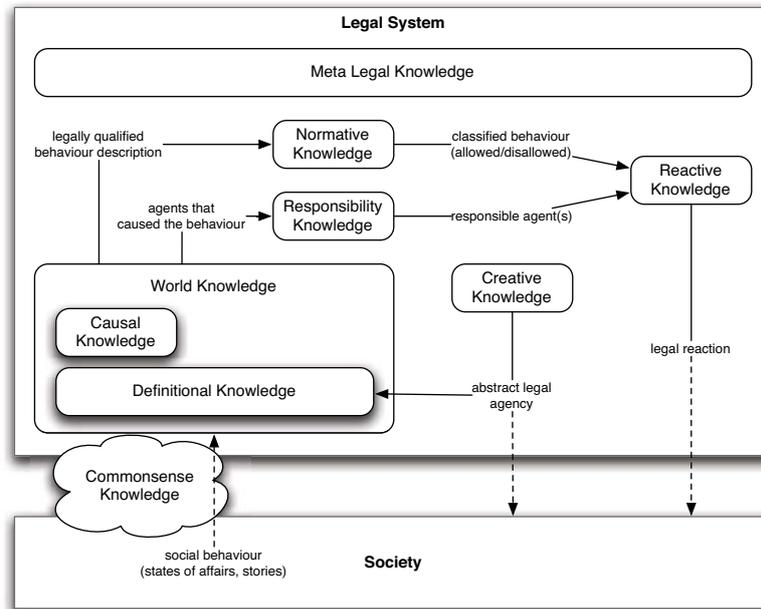


Figure 6.1: The Functional Ontology of Law

Normative knowledge reflects the regulatory nature of law, and has two functions: prescribing behaviour, and defining a standard of comparison for social reality. A norm expresses an idealisation: what ought to be the case (or to happen) according to the will of the agent that created the norm (Kelsen, 1991). *Meta-legal knowledge* governs relations between norms, and is applied when solving normative conflicts (Elhag et al., 1999), and determining the *preference ordering* of norms (Boer, 2009). Legal principles such as *lex specialis*, *lex superior* and *lex posterior* are types of meta-legal knowledge.

Because law governs the behaviour of agents in the world, it must contain some description of this behaviour. The separation between the knowledge used to describe the world and the normative knowledge was first explicitly proposed in Breuker (1990), where this category was called *world knowledge*, or the *Legal Abstract Model* (LAM). This model functions as an epistemological interface between the legal system and social reality. World knowledge consists of *definitional knowledge* and *causal knowledge* and defines legal concepts and (causality) principles as abstractions of commonsense concepts. The knowledge needed to attribute responsibility given the legal (causal) interpretation of some story or state of affairs, is characterised as *responsibility knowledge*.

The functional ontology of law describes the path along which a formal account travels as it is under consideration of the legal system (see Figure 6.1). When an account 'enters' the legal system, it is enriched with legal vocabulary, i.e. it is interpreted and legally *qualified*. The causal relations between events described in the account are identified. This interpreted causal account is then considered for norm violations, using some body of normative knowledge (rules, legislation, precedents), and liability is determined for agents

(in)directly responsible for the violation or some harm. A reactive measure is then taken on the basis of the outcome of this process.

### 6.1.2 Purpose

The FOLaw approach is reflected in various systems, such as the TRACS system on traffic regulations of den Haan (1992, 1996), the ON-LINE legal information server of (Valente, 1995; Valente and Breuker, 1999), and the MILE system for legal assessment in the ship insurance domain developed within the CLIME project (Winkels et al., 1998, 2002).<sup>7</sup> In these systems, definitional and normative knowledge were most prominently defined. In particular, the world knowledge of the MILE system was represented as a domain ontology comprising 3377 concepts and 11897 relations (Boer et al., 2001). It covered about 15% of the regulations for ship classification of Bureau Veritas, a large insurance firm in France, and all of the MARPOL I and II regulations for maritime pollution.<sup>8</sup> In all, the ontology had 8289 references to these legal sources. The CLIME ontology was used by the MILE system for *conceptual retrieval* on the body of regulations. Queries could be formulated in terms of concepts and relations of the ontology. This set was then expanded using a finite state automaton, and references to the regulations from this result set were presented to the user. The rationale for a particular result was explained as a minimal path from the concepts in the result set to the query set. Later, the ontology was reused in a *knowledge desktop environment* that provided workflow directed document management support for insurance surveyors (Jansweijer et al., 2001).

The way in which the CLIME ontology improved the accessibility of the Bureau Veritas regulations was a typical *knowledge management* use. For instance, the *legal encoding tools* (LET) editor that was used to build the ontology offered typical management functionality, such as automatic indexing and concept extraction (Boer et al., 2001).<sup>9</sup> The ontology itself was relatively lightweight; the language in which it was expressed supported a fixed set of five relations and did not have formal semantics.<sup>10</sup> A similar perspective on ontology use is e.g. exemplified by the DALOS ontology of Agnoloni et al. (2009), the intellectual property rights ontology IPRonto (Delgado and García, 2003), and the ontology for the Iuriservice portal (Casellas et al., 2007).

However, the CLIME ontology was also the basis for incremental specification of *normative* knowledge. Concepts in the ontology were used directly in the representation of norms. This proved to be a significant bootstrapping mechanism, not just in development speed, but also as a backbone for calculating the exception structure of norms as given by the *lex specialis* principle (Winkels et al., 1999). The legal assessment engine of MILE depended on inferences over the domain knowledge in the CLIME ontology.

In a similar vein, the LKIF includes a core ontology (LKIF Core) that provides

---

<sup>7</sup>CLIME, Cooperative Legal Information Management and Explanation, Esprit Project EP25414

<sup>8</sup>See <http://www.bureauveritas.com> and [http://www.imo.org/Conventions/contents.asp?doc\\_id=678&topic\\_id=258](http://www.imo.org/Conventions/contents.asp?doc_id=678&topic_id=258).

<sup>9</sup>It was in fact experience in developing the LET that initiated the MetaLex initiative of Boer et al. (2002).

<sup>10</sup>The five relations of CLIME were subsumption, part of, observable, measurable, related to and their inverse.

a vocabulary and a set of standard definitions of concepts common to all legal fields (Breuker et al., 2007; Hoekstra et al., 2007, 2008). In Breuker et al. (2007); Hoekstra et al. (2007, 2008) we identified three main ways for a legal core ontology to support information exchange in the setting of LKIF (cf. Valente (2005)). First of all, the ontology can serve as a resource for special, legal *inference*. Secondly, the definitions of terms in the ontology can facilitate *knowledge acquisition*, a terminological framework can facilitate the *exchange* of knowledge across multiple knowledge bases, and lastly it can be a basis for semantic *annotation* of legal information sources.

**Resource for Special, Legal Inference** Typical and abstract legal concepts are often strongly interrelated and thereby provide the basis for computing equivalencies, or paraphrases, and implications. For instance, by representing an obligation as the opposite of a prohibition, a (legal) knowledge system can make inferences that capture the notion that they are each others' inverse. A prohibition leaves all options open – except the one that is forbidden – while an obligation is unavoidable when all its requirements, or conditions, are satisfied. Although this implicit knowledge is relevant when reasoning with norms and cases, it does not express the control knowledge of reasoning (as in problem solving methods), but merely elicits the implications of declarative definitions. Specialised legal inference can be based on definitions of concepts in an ontology: an inference engine can generate the implied consequences of explicit concept definitions.

A classical example of specialised inference using the definitions in an ontology and a (general) inference engine is temporal reasoning based on Allen (1984)'s ontology of time (Section 6.3.1). To enable special inference, terms should be highly interrelated and form a coherent *cluster* with little or no external dependencies (Hayes (1985), and Section 5.2). An example of such a cluster in the legal domain is that of the terms that denote deontic qualifications. Clusters of this type are usually found at very high levels of abstraction.

**Knowledge Acquisition Support** The classical use of both top and core ontologies in knowledge representation is as a means to support knowledge acquisition (Section 5.4). If well designed and explained, they provide an initial structure to which domain terms can be attached as subclasses. Inheritance of properties and other implicit knowledge can then be used to check not only consistency, but also the extra-logical quality of the ontology: whether what is derived (classes, properties) makes sense. The use of a core or top structure that has well tested and evaluated implications, makes it easier to check whether domain refinements are not only consistent, but also arrive at inferences that correspond to what the knowledge engineer or user holds to be valid. The knowledge acquisition support of ontologies is not restricted to just ontological or even terminological knowledge. For instance, the incremental specification of normative knowledge in the CLIME ontology is an example where an ontology aids in the knowledge acquisition of non terminological knowledge.

**Preventing Loss in Translation** A legal ontology can play an important role in the translation of existing legal knowledge bases to other representation formats. In particular when these knowledge bases are converted into LKIF

as the basis for articulate knowledge serving. Similar to a translation between different natural languages, a formal, ‘syntactic’ translation may clash with the semantics implied by the original knowledge representation. An ontology, as representation of the semantics of terms, allows us to keep track of the use of terms across multiple knowledge bases.

**Resource for Semantic Annotation** The semantic annotation of legal sources, such as regulations and jurisprudence is an important contribution to the accessibility and maintainability of these sources. First of all, an ontology can be a source for information retrieval, such as e.g. the CLIME ontology. Secondly, the status of an officially sanctioned legal text is primarily determined by its relation to other legal texts (Boer et al., 2004a). This status can be made explicit by expressing it using ontologically defined relations. In fact, these relations do not just hold between the texts themselves, but between the formal representations of their content as well (Klarman et al., 2008).

Besides the general requirements for knowledge representation ontologies outlined in the preceding chapters, the ontology is to contain a core set of definitions for describing specific legal terms. We have seen that law can be viewed as an instrument used by the legal and political system to identify and control situations and events in social interaction (Valente, 1995). By far the bulk of social situations, be it in our family life, at work, related to transport, property, crime, etc. is not described in specialised technical terms: their meaning is part of common sense. For instance, the conflicts and problems brought to court – legal cases – are initially described using common sense terms, and are gradually translated into legal technical terminology in the process of coming to a decision. For this legal qualification to be possible, the gap between the vocabulary of a case and legal terminology needs to be bridged (Winkels and de Bruijn, 1996). The possibility of legal qualification in general – by legal professionals – is a strong indication that the vocabularies of legal knowledge and common sense are not disjoint. Legal terminology can be reduced to the actual societal events and states governed by law. In other words, the basic categories of the LKIF ontology should reflect Valente’s view that legal world knowledge is an abstraction of common sense.

A third requirement for LKIF is given by the ideal of the Semantic Web to achieve understanding both between web services and between human users (Chapter 3). In fact, a commonsense perspective is also applicable to any serious endeavour towards a Semantic Web. As the the web is about the most diverse information source we know today, a common sense oriented ontology would certainly be an important first step to more uniform web-based knowledge representation. Conversely, as hinted at in the introduction, the distribution of legal information across various strongly interconnected sources is a demanding use case for semantic web technology (Hoekstra, 2008). An important requirement is therefore that the LKIF core ontology should be represented using the DL profile of OWL.

The following sections describe the theoretical and methodological framework against which the LKIF core ontology has been developed. Section 6.2.1 discusses the perspective used in its construction in relation to five other ontologies. The methodology used to construct the LKIF ontology is discussed in

Section 5.2. Section 6.3 introduces the modules and most important concepts of LKIF Core.

## 6.2 Developing a Core Ontology

Given a commonsense perspective, it is expected that (at least parts of) existing ontologies would be reusable. Either as a source of inspiration or for the purpose of full formal reuse of definitions. This would hold in particular for top ontologies that include legal terms, as for instance listed in Casanovas et al. (2006). Unfortunately, it turned out that the amount of reuse and inspiration was rather limited. Not only do existing ontologies diverge on the approach, coverage and knowledge representation language used; those ontologies that do claim a common sense or similar perspective differ in their conception as to what such a perspective *means*.

### 6.2.1 Ontology Integration

This section evaluates several of the foundational and core ontologies introduced in Section 5.4.1 with respect to a potential contribution to LKIF Core. The main requirement is their suitability to enable the primary roles of the LKIF ontology outlined in the previous section. We pay specific attention to their definition of commonsense and legal terms, and possibilities for safe reuse (Section 5.4.2).

#### Suggested Upper Merged Ontology

The SUMO ontology of Niles and Pease (2001) brings together insights from engineering, philosophy, and information science. It provides definitions for general purpose terms, is intended as a unifying framework for more specific domain level ontologies. As a starting point for a legal core ontology SUMO has several drawbacks. First of all, it does not readily provide definitions of terms relevant to the legal field – e.g. its coverage of mental and social entities is limited. Because of the way in which SUMO is constructed, it has a bias towards more abstract and theoretical insights coming from engineering and philosophy. Although it is non-revisionist, as in e.g. the distinction between objects and processes, it does not have a real commonsense basis. Furthermore, as discussed in Section 5.4.1 SUMO is a *foundational ontology* and uses meta modelling, such as in the definitions of classes, binary relations and sets. As SUMO is represented in the expressive language KIF, and more recently CommonLogic, this practice is not fundamentally problematic. However, it means that safe reuse of SUMO definitions in the construction of an OWL DL ontology is not possible.

#### Descriptive Ontology for Linguistic and Cognitive Engineering

DOLCE is part of the WonderWeb library of foundational ontologies, cf. Masolo et al. (2003); Gangemi et al. (2002). It was meant as a reference point for the other ontologies in the library, to make explicit the differing rationale and alternatives underlying the ontological choices of each ontology. Rather than a

coherent set of ontological commitments, it captures a range of alternatives. This way, the library would form a network of different but systematically related ontology modules. The relation between an ontology, available in the library, and the DOLCE ontology expresses its ontological commitment to particular ontological options. DOLCE was therefore never presented as the foundational ontology it is currently regarded as, but it has been successfully used as such in a large number of projects.

DOLCE is very much an ontology in the philosophical tradition, and differs from the knowledge representation perspective in two significant ways. Firstly, its perspective is philosophical with respect to its *content*, i.e. it is aimed to directly represent reality. And secondly, it is subject to the epistemological promiscuity of philosophical ontology because it is rather an extension of the knowledge representation formalism at the *ontological level* (Guarino, 1994), than a model expressed using that formalism. The meta-level character of DOLCE means that the ontology is not a representation of knowledge, but of the terms used to describe knowledge – in the same way that the constructs of OWL are. Like SUMO, DOLCE was originally specified in first order logic and the highly expressive KIF language. Its OWL DL representation (DOLCE-Lite) is more restrictive, e.g. it does not consider temporal indexing and relation composition.

The DOLCE ontology is *descriptive*, and is based on the stance that “the *surface structure* of natural language and human cognition”<sup>11</sup> is ontologically relevant. It is argued that this perspective results in an ontology that captures cognitive artefacts more or less depending on human perception, cultural imprints and social conventions, and *not* deep philosophical insights of Ontology. DOLCE thus claims an approach that fits more with a commonsense perspective than the science perspective of SUMO. However, the suggestion that this surface structure has any bearing on common sense is not based on evidence. Rather, the methodological commitment to the *surface* structure of language and cognition almost inevitably resulted in an intricate framework of theoretical notions needed to encompass the idiosyncrasies of human language use. Alternatively, rather than constructing an ontology by studying reality through the kinds of categories implicit in natural language, a pragmatic approach based more directly on the conceptualisation of reality we use and live by in our daily routine is more appropriate (Breuker and Hoekstra, 2004a; Hoekstra et al., 2008, and Section 6.2.2).

### Core Legal Ontology

Over the years, DOLCE has been extended in several ways. DOLCE+ is the extension of DOLCE with a theory on descriptions and situations (also called D&S, Gangemi and Mika (2003)). CLO, the Core Legal Ontology (Gangemi et al., 2005) extends DOLCE+ even further and defines legal concepts and relations based on its formal properties. CLO was designed to support both the definition of domain ontologies, a juridical Wordnet, and the design of legal decision support systems. To a large extent these goals correspond with the requirements of the LKIF ontology.

CLO conceives the legal world as a *description* of social reality, an ideal view of the behaviour of a social group. It builds on the D&S distinction between

<sup>11</sup>Emphasis by the authors, Masolo et al. (2003)

*descriptions*, and *situations*. Examples of legal descriptions, or *conceptualisations*, are the *contents* of laws, norms, and crime types. These descriptions constrain legal situations, i.e. legal *facts* of *cases*, such as legal, relevant non-legal and juridical states of affairs. Every legal description *classifies* a state of affairs. More precisely, a legal description is the reification of a theory that formalises the content of a norm, or a bundle of norms. A legal case is the reification of a state of affairs that is a logical model of that theory. A description is satisfied by a situation when at least some entity in the situation is classified by at least some concept in the description. Classification in CLO is thus not DL classification, and it is unclear as to what extent the two interpretations are compatible.

As described earlier, the legal system as description, or rather *prescription*, of reality is not new, cf. Valente (1995, others). However, the CLO distinction between descriptions and situations is rather one dimensional in that it does not commit to an ontological view of the *kinds* of descriptions involved. In line with the DOLCE ontology, it confounds the distinction between representation and the *represented* with representation and *reality*. Although it introduces new levels of abstraction by reification, it does not provide ontological categories that can be used to describe the knowledge at these levels. In a language that itself can be conceived as providing the means to construct descriptions of reality (situations), such as OWL DL, it is unclear what the epistemological status of the classes ‘description’ and ‘situation’ *themselves* is. For instance, what is the difference between an individual description being classified-in-the-OWL-sense as some description class, and a situation class being classified-in-the-CLO-sense by that same description?

As CLO relies on a subset of DOLCE for the definition of *elements* of situations, it is subject to the same criticism with respect to its commonsense perspective. Moreover, the lack of ontological commitment at the level of descriptions undermines its suitability for knowledge acquisition support in a legal setting as well. Although for sure a norm can be described as some description of a situation, it is not the norm-as-description that uniquely characterises what a norm *is*. This holds especially for less obvious ‘descriptions’ (in CLO terms), as e.g. damage or right of way.

## CYC

CYC is a huge unified ontology of commonsense concepts (Lenat et al., 1990; Lenat, 1995). Although the project has started as early as 1984, its general set-up corresponds to that of later large scale ontology efforts. The main motivation for the Cyc Project was that all software programs would benefit from the availability of commonsense knowledge. This idea is not new, and was acknowledged in the early years of AI: “A program has common sense if it automatically deduces for itself a sufficiently wide class of immediate consequences of anything it is told and what it already knows” (McCarthy, 1959, p.2).

The idea is that, when enough commonsense knowledge is represented, and a certain threshold is reached, a quantum-leap (“The Singularity”)<sup>12</sup> would enable CYC to expand its knowledge through guided learning (as a human child would). This theory is in line with Minsky’s ideas about how computers can become intelligent beings: add enormous amounts of commonsense know-

<sup>12</sup>Indeed, CYC is a much-hyped project, and has received a lot of criticism because of it.

ledge (Minsky, 1982, 1984). This basic knowledge about the workings of the world would finally allow you to send the kid to school. With currently over 300K concepts, the knowledge base seems well under way in reaching this threshold, however we still have to see the first results.<sup>13</sup>

The upper part of the CYC ontology is claimed to express a commonsense view and indeed it is more concrete than either SUMO or DOLCE. On the other hand, from a methodological point of view, the CYC approach is not very satisfactory either. Technically, CYC qualifies rather as a terminological knowledge base than as an ontology proper.<sup>14</sup> Instead of a meticulous study of the actual workings of the world, as in SUMO, or the surface structure of language and cognition, as in DOLCE, it seems the approach followed is to have a bunch of knowledge engineers simply put everything they know into the CYCL formalism.<sup>15</sup> This procedure results in a large portion of the knowledge base being decidedly non-ontological, but rather context dependent *frameworks* (see Section 5.5.2).

Furthermore, CYC suffers from two more technical impediments for reuse. Firstly, like SUMO and DOLCE, it is specified in the very expressive CYCL representation language, which is based on first-order predicate calculus. Recently a part of the publicly available OpenCYC effort has been made available in OWL Full, but again, this does not cover the full semantics of the ontology.<sup>16</sup> CYCL admits meta modelling, which indeed is used liberally throughout the ontology – even more so than in DOLCE and SUMO. Secondly, the sheer size of the knowledge base – as with most unified ontologies – introduces significant reasoning overhead for even the simplest tasks. As such, CYC seems more suitable for direct inclusion in a knowledge based system than as a conceptual coat rack for ontology development on the Semantic Web.

### 6.2.2 Ontology Reuse

Thus far, the ontologies we reviewed do not appear to meet the requirements for the top structure of a legal core ontology. Although in the past few years the ontologies underwent changes and extensions, these results are in line with the outcome of an earlier review (Breuker and Hoekstra, 2004a). Firstly, the ontologies are specified at multiple (meta) levels of abstraction, using very expressive languages, which limits possibilities for safe reuse.

Furthermore, in all three ontologies those concepts that are of relevance to law are either scarce and under specified, or overly theoretical. In particular, our requirement that a legal core ontology should be built on a commonsense conception of reality is not met. Where a commonsense perspective is claimed, it is not motivated, explained or substantiated. The common sense of CYC is in fact *common knowledge*, or rather “human consensus reality knowledge” (Lenat et al., 1990, p. 33), i.e. that knowledge of things most humans will concede to exist in reality. In contrast, the DOLCE and SUMO ontologies do not commit

<sup>13</sup>The online game FACTory for teaching CYC is online at <http://game.cyc.com>.

<sup>14</sup>A random concept search just returned ‘Laurel Goodwin’, the acclaimed American actress who played the memorable role of Yeoman J.M. Colt in the 1966 pilot of Star Trek. See <http://sw.cyc.com/2006/07/27/cyc/LaurelGoodwin>.

<sup>15</sup>See <http://www.cyc.com/cycdoc/ref/cycl-syntax.html>

<sup>16</sup>See <http://www.opencyc.org> for an online browser and <http://www.cyc.com/2004/06/04/cyc> for the OWL Full version

to the ontological relevance of human consensus, but rather aim to provide an ontological grounding for *all* knowledge. They do this in two distinct ways. SUMO is based on scientific knowledge of objective reality, and for DOLCE the way in which we apply and consciously report our knowledge is ontologically relevant.

Both approaches are generic enough to be the basis for an ontology that describes our common sense, in the sense of common knowledge. However, in our view, common sense refers not to some common body of knowledge of reality, but rather to the commonality of the scope and detail of that knowledge as it manifests itself in individual persons: it is a *level of description*, much akin to the *basic level* of Lakoff (1987) (see Section 5.2). Characteristic of this level is not just the *kind* of things we commonly know, but more importantly, the way in which this knowledge is *structured*. But for the last requirement, a philosophical approach would be perfectly adequate. However, a commonsense ontology should not be specified in highly specific, theoretical jargon, but should rather have a commonsense structure of its own.

### LRI Core

The review in Breuker and Hoekstra (2004a) motivated a decision to develop a legal core ontology to support the development of ontologies for criminal law in various European countries, as part of the e-Court project.<sup>17</sup> This ontology, *LRI Core*, was developed with a set of design goals similar to that of LKIF Core, cf. Breuker and Hoekstra (2004c,a). What sets LRI Core apart from other ontology efforts is that its definitions were aimed to be verifiable through empirical research on how humans relate concepts in actual understanding of the world; and not about revisionist views of how we should view the world as it actually *is* (as e.g. in correct theories of the physical world) or as it makes up a parsimonious view on reality (as e.g. in philosophical views on the main categories of description).<sup>18</sup> This kind of empirical evidence can range from cluster analysis of semantic distance between terms, to neuro-psychological evidence.

Central to this effort is the view that common sense is rooted in a conceptualisation that is at its heart the result of our evolution. This conceptualisation is developed in order to deal with a dynamic and potentially dangerous environment. Our capacity to move, perceive and interact with reality has led to increasingly complex cognitive representations. These representations range from hard-wired abstraction in our perception system, such as the ability to perceive straight lines and angles at a mere 2 neurones away from the retina, via the inborn syntheses of perceptual input into basic properties, to – eventually – the representations accessible to conscious thought.

This range of increasingly abstract and complex representations of reality defines an axis that indexes organisms in successive stages of evolutionary development, e.g. from viruses and bacteria to multi-cell organisms, insects, mammals, primates and finally homo sapiens. In other words, the compet-

<sup>17</sup>Electronic Court: Judicial IT-based Management. e-Court was a European 5th Framework project, IST 2000-28199.

<sup>18</sup>For instance the classical distinction between abstract and concrete concepts in philosophy does not fit well with human understanding of the dynamics of the world in terms of physical causation, intention and life. The commonsense explanation of an event may involve all three categories.

encies of our genetic ancestors give insight in what the roots of our common sense are. These roots may well be hidden too deep to be accessible to conscious thought and introspection. Nonetheless, on the basis of insights in cognitive science we can make several basic distinctions that go well beyond a mere hunch. As most are of relevance to the LKIF Core ontology as well, they are briefly outlined below.

**A Cognitive Science Perspective** Given that the physical environment is relatively stable – the notable exceptions being day-night cycles, changing weather conditions and the occasional geological perturbation – the perceptual apparatus of most organisms is tuned to register the slightest change occurring against this stable canvas. Of particular relevance is the ability to be aware of changes induced by other organisms. Firstly, the presence of another organism may present an opportunity for *reproduction* and sustaining *metabolism* (i.e. by eating). And secondly, their presence may signify a direct threat to an organisms' existence. The result is a prominent distinction in cognition between 'background' and 'foreground' (Hobbs, 1995). In general our awareness is directed to the *discontinuity* of change rather than spatial arrangements of objects or historical continuity. The ability to perceive stability is enabled by episodic memory, though it requires some conception of the physical constraints underlying this stability. An example is the general rule that physical objects keep their position unless subjected to a change in the exertion of force. Changes occur against the canvas of temporal and spatial positions, and the speed of a change is indicative of whether it becomes foreground or remains in the background.

In LRI Core, the view that knowledge serves to interpret occurrences in the world was reflected by a contrast between concepts on the one hand, and individuals and their occurrence (instance) on the other. Furthermore, the cognitively primary distinction between static and dynamic elements in the (physical) world is reflected by differentiating objects and processes. Objects have extensions in space where processes have extensions in time, but are contained by objects. Processes reflect a causal explanation of change. The notions of space and time do not just play a role as the extension of objects and processes but can indicate positions as spatio-temporal referents. The spatio-temporal position of objects is not inherent: a change in position does not constitute a change *in* an object.

Very recent – at least in evolutionary terms – mammals developed the ability to attribute intentions to other animals. Because of the intentional interpretation of behaviour, change is no longer private to physical causation but to the *actions* of other *agents* as well. Actions are always intentional “under some description” (Davidson, 2001, ch.3): performing an action comes down to the initiation of processes that bring about some intended change or state.<sup>19</sup> Paradoxically enough, taking into account the mental state of other animals precedes the ability to consciously reflect on our *own* mind. And although this was long thought to be a skill exclusive to humans, it has been shown that our next of kin – chimpanzees, bonobo's – have self awareness as well.

<sup>19</sup>Note that this allows the same events to be not intentional under some other description. However, if we take the events to constitute an *action*, this is a description that presupposes intentionality.

Because mental representation of other mental representations is a fairly recent accomplishment, the models we construct reuse many parts of the conceptualisation originally developed to deal with physical reality. We think and speak of mental *processes*<sup>20</sup> and mental *objects* in the same terms we use to talk about the physical world. In other words, these terms are metaphors of similar physical notions. Some, i.e. Lakoff and Núñez (2000), even argue that the same mechanism is used to construct the highly abstract notions of mathematics.

Social awareness and self awareness are the most important prerequisites for complex social behaviour. First of all, they enhance the predictability of our environment by allowing us to take the possible intentions of other agents into consideration for planning and control. Secondly, it allows us to share plans with other agents. Repeated execution of such co-operative plans can render them institutionalised by a community. Plans make extensive use of *roles* to specify expected or default behaviour. The ability to play a certain role expresses a (recognised) competence, sometimes acknowledged as a social 'position'. In LRI Core, roles played a central part in the construction of social structures.

The LRI Core ontology distinguished four 'worlds':

1. A *physical world*, divided by processes and objects, each containing matter and energy.
2. A *mental world*, containing mental objects and mental processes. The mental world is connected to the physical world through actions, which translate an intention into some physical change.
3. A *social world*, built from mental objects such as roles
4. An *abstract world*, which contains only formal, mathematical objects.

Because of its grounding in cognitive science and its explicit common sense perspective, the characteristics of the LRI Core ontology are relatively similar to those intended for the LKIF ontology – especially in comparison to the other ontologies we discussed so far. And it is indeed true that it can in many ways be seen as the direct precursor of the LKIF ontology.<sup>21</sup>

Nonetheless, the LRI Core ontology falls short in several important respects. Though it is specified in OWL DL, most concepts in the ontology are under specified. They are defined by subsumption relations, and are only sparsely characterised using more complex class restriction. Furthermore, the distinction between the different worlds in LRI Core is fairly absolute, and no theory is provided as to how they are connected and interact. Thirdly, apart from a relatively well-developed characterisation of roles, LRI Core is relatively under-developed with respect to the mental world. For instance, it emphasises physical objects, processes, energy and substance while remaining rather sketchy with respect to their mental counterparts. In part the limitations of LRI Core can be ascribed to an unprincipled top down methodology (see Section 5.3).

Concluding, although the perspective and main distinctions of LRI Core were used as inspiration in the construction of the LKIF ontology, it is not

<sup>20</sup>Also: mental actions, e.g. in (trying) to control one's thoughts.

<sup>21</sup>This is not really surprising as there exists an overlap between the developers of LRI Core and LKIF Core.

#	Importance	Abstractness	Legal Relevance
1	Law	Deontic operator	Civil law
2	Right	Law	Law
3	Jurisdiction	Norm	Legal consequence
4	Permission	Obligative Right	Legislation
5	Prohibition	Permissive Right	Obligation
6	Rule	Power	Right
7	Sanction	Right	Authority
8	Violation	Rule	Deontic operator
9	Power	Time	Duty
10	Duty	Anancastic Rule	Jurisdiction
11	Legal Position	Existential Initiation	Legal Fact
12	Norm	Existential Termination	Legal Person
13	Obligation	Potestative Right	Legal Position
14	Permissive Right	Productive Char.	Legal Procedure
15	Argument	Absolute Obl. Right	Liability

Table 6.1: Fifteen highest scoring terms for *importance*, *abstractness*, and *legal relevance*.

simply a specialisation of LRI Core. Not only is it built from the ground up, the methodology by which it is constructed forces a broader, more concrete, and more rigorous definition of concepts and relations. First, the scope of the ontology was determined by selecting a core set of basic concepts. These concepts were organised in modules, and formed the basis for a middle-out construction of the ontology.

### 6.2.3 Scope

The LKIF Core ontology is to provide a core set of definitions to be used across the legal domain. Not only should these definitions be relevant to law, they should exist at a level that allows for a more specific interpretation of terms in sub domains of the field such as criminal or private law. Furthermore, a focus on overly theoretical and abstract concepts impedes the connection to common sense. In other words, we are looking for the *basic concepts* of law (Lakoff, 1987, and Section 5.3.1) that allow us to construct the ontology in a middle-out fashion (Uschold and King, 1995, and Section 5.3).

As discussed in Section 5.3.1, what concepts count as basic depends on the expertise present in a community of practice. Where ‘manslaughter’ may be basic to legal professionals, the average citizen is unlikely to be able to explain how it is different from murder. To obtain an appropriate level of description, the basic concepts in the ontology should reflect the heterogeneity of its prospective users. In Breuker et al. (2007) we identified three main groups of users: *citizens*, *legal professionals* and *legal scholars*. Although legal professionals use the legal vocabulary in a far more precise and careful way than laymen, for most of these terms the extent of common understanding is sufficient to treat their usage as analogous (Lame, 2004). Nonetheless, a number of basic terms, such as ‘liability’ and ‘legal fact’, have a specific technical legal meaning.

Several representatives from every partner in the Estrella consortium, covering each of the three user groups were asked to supply their Top 20 of legal concepts. These lists were combined with terms frequently occurring in literature such as jurisprudence and legal text-books. The resulting list comprised a

#	Legal vs. Commonsense	Common vs. Specific
1	Perpetual injunction	Absolute right
2	Cantonal Court	Fact
3	Elements of claim	Jurisdiction
4	Provisional rule	Law
5	Solicitor	Lawyer
6	Interlocutory injunction	Obligation
7	Other directed norm	Permission
...	...	...
242	Critical question	Narcotics
243	Dialogue	Pensioner
244	Dialogue type	Veteran
245	Family	Deformation
246	Criteria	Genocide
247	Critical discussion	Enthymemes
248	Mafia	Mafia

Table 6.2: Top and bottom seven terms for *legal vs. common sense* and *common vs. specific legal term*

grand total of about 250 terms. Users were asked to assess each term from this list on five scales:

- the perceived level of *abstraction*,
- the perceived *relevance* for the legal domain,
- the degree to which a term is *legal* rather than *common-sense*,
- the degree to which a term is a *common legal term* (as opposed to a term that is specific for some sub-domain of law),
- and the degree to which the expert thinks this term should be *included* in the ontology.

The resulting scores were used to select an initial set of 50 terms plus several reused from other ontologies (see Tables 6.1 and 6.2), and formed the basis for the identification of clusters and the development of the LKIF Core ontology. In the end, several technical terms were included in the ontology because they capture an abstract meaning central to law (Sartor, 2006). Furthermore, such terms provide structure to the relation between more generally understood legal terms.

### 6.3 Ontology Modules

The preceding sections introduce the requirements for LKIF Core as a knowledge representation ontology for the legal domain. It is based on insights from cognitive science and uses the well established middle-out methodology. With these considerations in mind, the LKIF ontology was initially designed as a collection of eight ontology modules: *expression, norm, process, action, role, place, time* and *mereology*, cf. Breuker et al. (2006). This collection was later extended with a top ontology and two more ontology modules (*legal\_action, legal\_role*), see Figure 6.2 (Hoekstra et al., 2007; Breuker et al., 2007).<sup>22</sup> Each of these mod-

<sup>22</sup>Version 1.0 of the ontology also included two frameworks (*time\_modification* and *rules*) but these are not discussed here.

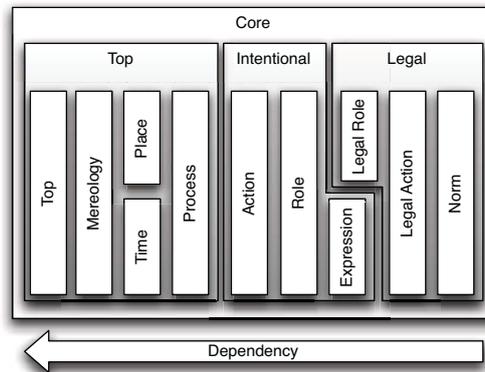


Figure 6.2: Dependencies between LKIF Core modules.

ules contains a relatively independent cluster of concepts that was represented using OWL DL in a middle-out fashion: for each cluster the most central concepts were represented first.<sup>23</sup>

The ontology modules are organised in three layers: the *top* level (Section 6.3.1), the *intentional* level (Section 6.3.2) and the *legal* level (Section 6.3.3). These layers correspond to the different *stances* one can adopt towards a domain of discourse, and are inspired by the work of Dennett (1987) and Searle (1995). Dennett identified three stances we can adopt for explaining phenomena in the world: the *physical* stance, used for explaining behaviour in terms of the laws of physics, the *design* stance, which assumes a system will behave according to its design, and the *intentional* stance, which can be adopted to explain the behaviour of rational agents, with beliefs, desires and intentions. The first two correspond roughly to the top level of LKIF Core, where the intentional stance is captured by the intentional level:

“that feature of representations by which they are *about* something or *directed at* something. Beliefs and desires are intentional in this sense because to have a belief or desire we have to believe that such and such is the case or desire that such and such be the case. Intentionality, so defined, has no special connection with intending. Intending, for example, to go to the movies is just one kind of intentionality among others.”

(Searle, 1995, p.7)

The LKIF ontology thus adds a *legal* layer, containing concepts that are only sensible from a legal perspective. Accordingly, we represent social and legal concepts as social constructs separate from the physical entities having such imposition, e.g. persons are said to play roles and objects fulfil functions (Searle, 1995).

This distinction does not hold for the layers as a whole: they should not be confused with stratified meta levels. Each layer introduces a straightforward

<sup>23</sup>The ontology was developed using TopBraid Composer (<http://www.topbraidcomposer.com>) and Protégé 4.0a (<http://protege.stanford.edu>).

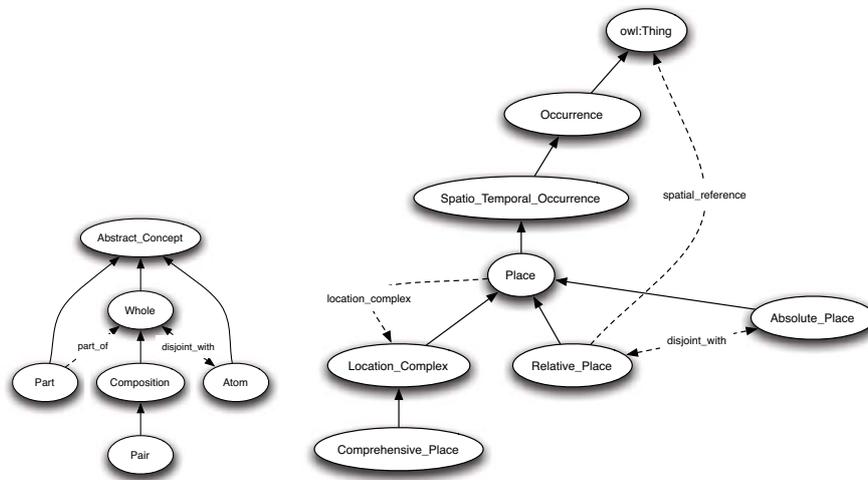


Figure 6.3: Place and Mereology related concepts.

extension of existing definitions. At each level, concepts are expressed in terms of concepts defined at a higher level of abstraction, adding new organising structures (such as properties) where necessary. However, module extension does not follow the requirement of safe reuse (Section 5.4.2). As regards its ontological commitment the ontology should be regarded as a *whole*. This means that safety of reuse of the ontology by third parties can only be assessed with respect to the combination of all ontology modules. This methodology ensures a modular set-up that improves reusability and allows extensions to commit to the ontology at one of the levels, without compromising compatibility with extensions that commit to a different level.

The following sections give a concise overview of the modules of LKIF Core and their principal concepts. They do not provide exact OWL DL definitions of these concepts but rather focus on the necessary ingredients for a legal core ontology. Several principal concept definitions are discussed in detail in the next chapter.

### 6.3.1 First Things First: The top-level

The description of any legally relevant fact, event or situation requires a basic conceptualisation of the context in which these occur: the backdrop, or canvas, that is the physical world. Fundamental notions such as location, time, mereology and change are indispensable in a description of even the simplest legal account. The top level clusters of the ontology provide (primitive) definitions of these notions, which are consequently used to define more intentional and legal concepts in other modules. The most general categories of the LKIF ontology are based on the distinction between ‘worlds’ of LRI Core. We distinguish between mental, physical and abstract entities, and occurrences. Mental entities reside in the human mind, and only have a temporal extension. Physical entities exist independent of human experience, and have a spatial extension as well. Although these categories are superimposed on the concepts in the top

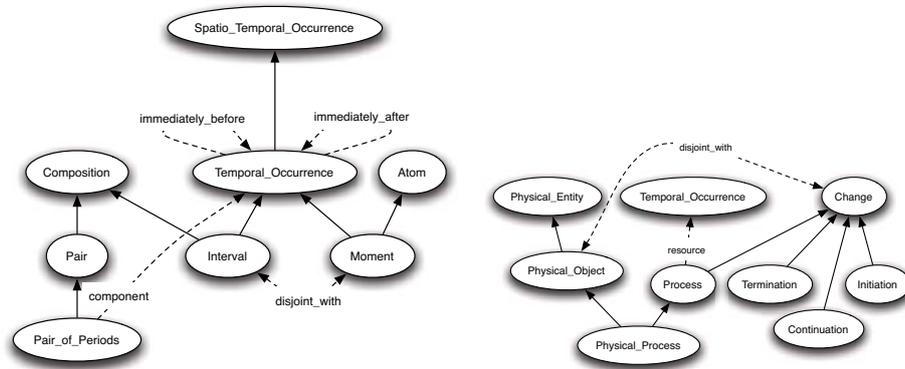


Figure 6.4: Concepts related to time and change.

level clusters, they were not directive in their design.

Mereological relations allow us to define parts and wholes; they can be used to express a systems-oriented view on concepts. Examples are functional decompositions, and containment characteristic for many frameworks (Figure 6.3). Mereology lies at the basis of definitions for *places* and moments and intervals in *time*. The ontology for places is based on the work of Donnelly (2005) and adopts the Newtonian distinction between *relative* and *absolute* places. A relative place is defined by reference to some thing; absolute places are part of absolute space and have fixed spatial relations with other absolute places. A *Location\_Complex* is a set of places that share a common reference location, e.g. the locations of all furniture in a room. See figure 6.3 for an overview of concepts defined in the place module. Of the properties defined in this module, *meet* is the most basic as it is used to define many other properties such as *abut*, *cover*, *coincide* etc. See Breuker et al. (2007); Donnelly (2005) for a more in depth discussion of these and other relations. Currently this module does not define classes and properties that express direction and orientation.

Similar relations can be used to capture notions of time and duration. We adopt the theory of time of Allen (1984); Allen and Ferguson (1994), and distinguish between the basic concepts of *Interval* and *Moment*. Intervals have an extent (duration) and can contain other intervals and moments. Moments are points in time, they are atomic and do not have a duration or contain other temporal occurrences (see figure 6.4). Relations between these temporal occurrences can be used to express a timeline. Allen introduced the *meet* relation to define two immediately adjacent temporal occurrences. To discern between the temporal *meets* relation and its spatial counterpart (Donnelly, 2005), this relation is called *immediately\_before*. Where the spatial relation is unrestricted with respect to direction, the temporal *meet* relation is directed and asymmetric. It is used to define other temporal relations such as *before*, *after*, *during*. Locations and temporal entities are used to define the extension of mental and physical entities, e.g. the time when you had a thought, the location where you parked your car. They are *occurrences* and do not have extensions themselves, they *are* extensions.

With these classes and properties in hand we introduce concepts of (invol-

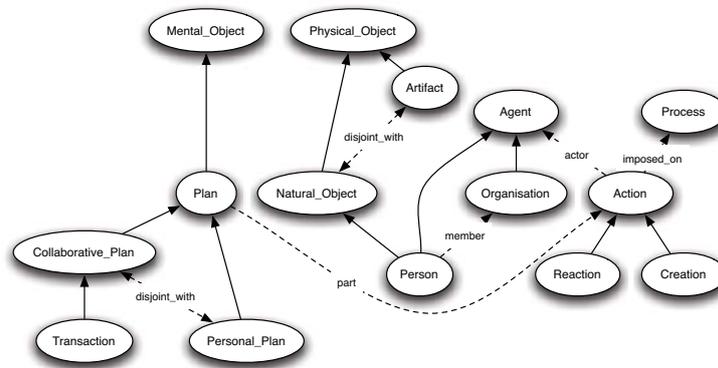


Figure 6.5: Actions, agents and organisations.

untary) change. The process ontology relies on descriptions of time and place for the representation of duration and location of changes. A Change is defined as a difference between the situation before and after the change. It can be a functionally coherent aggregate of one or more other changes. More specifically, we distinguish between Initiation, Continuation and Termination changes.

Changes that occur according to a certain recipe or procedure, i.e. changes that follow from causal necessity are Processes. They thus introduce causal propagation and are said to *explain* the occurrence of change. Processes in LKIF Core are similar to the fluents of event calculus (van Lambalgen and Hamm, 2005). However, the ontology does not commit to a particular theory of causation and we consider the perspective generic enough to enable the definition of various disparate conceptions of causal relations. Contrary to changes, processes are both spatially and temporally restricted. They extend through time – they have duration – and are located at some temporal and spatial position. We furthermore distinguish Physical\_Processes which operate on Physical\_Objects.

### 6.3.2 The Intentional Level

Legal reasoning is based on a common sense understanding that allows the prediction and explanation of intelligent behaviour. After all, it is only the behaviour of rational agents that is governed by law. The modules at the intentional level include concepts and relations necessary for describing this behaviour: Actions undertaken by Agents in a particular Role. Furthermore, it introduces concepts for describing the mental state of these agents, *subjective entities* such as their Intention or Belief, but also communication between agents by means of Expressions (Searle, 1995, and Section 7.3.2).

The class of agents is defined as the set of things which have intention and can be the actor of an action: they may perform the action and are potentially liable for any effects caused by the action (see Figure 6.5). An action is 'intentional under some description' (Davidson, 2001), they constitute an intentional interpretation of a process that reflects the changes brought about by some agent in realising his intentions. Agents are the medium of an action's intended outcome: actions are always intentional. The intention held by the agent usu-

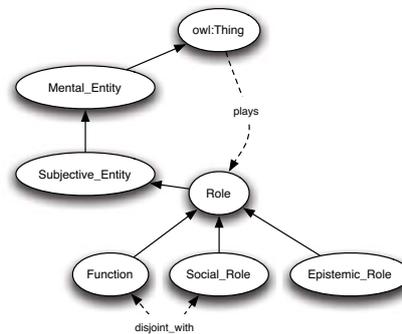


Figure 6.6: Roles.

ally bears with it some expectation that the intended outcome will be brought about: the agent believes in this expectation. The actions an agent is expected or allowed to perform are constrained by the *competence* of the agent, sometimes expressed as *roles* assigned to the agent. Because actions coincide with processes, they can play a role in causal propagation, allowing us to reason backwards from effect to agent (see Section 7.5). Actions can be creative in that they initiate the coming into existence of some thing, or the converse terminate its existence. Also, actions are often a direct *reaction* to some other action.

The LKIF Core ontology distinguishes between *Persons*, individual agents such as “Joost Breuker” and “Pope Benedict XVI”, and *Organisations*, aggregates of other organisations or persons which acts ‘as one’, such as the “Dutch Government” and the “Sceptics Society”. *Artefacts* are physical objects designed for a specific purpose, i.e. to perform some *Function* as instrument in a specific set of actions such as “Hammer”. *Persons* are physical objects as well, but are not designed (though some might hold the contrary) and are subsumed under the class *Natural\_Objects*. Note that natural objects can function as tools or weapons as well, but are not *designed* for that specific purpose.

The notion of roles played an important part in recent discussions on ontology (Steimann, 2000; Guarino and Welty, 2002; Masolo et al., 2004; Loebe, 2007). A *Role* is a *Subjective\_Entity* that specifies standard or required properties and behaviour of the entities playing the role (see Figure 6.6). Roles not only allow us to categorise objects according to their prototypical use and behaviour, they also provide the means for categorising the behaviour of other agents. They are a necessary part of making sense of the social world and allow for describing social organisation, prescribe behaviour of an agent within a particular context, and recognise deviations from ‘correct’ or normal behaviour. Indeed, roles and actions are closely related concepts: a role defines some set of actions that can be performed by an agent, but is conversely defined by those actions. The role module captures the roles and functions that can be played and held by agents and artefacts respectively, and focuses on *social* roles, rather than traditional thematic or relational roles.

A consequence of the prescriptive nature of roles is that agents connect expectations of behaviour to other agents: intentions and expectations can be

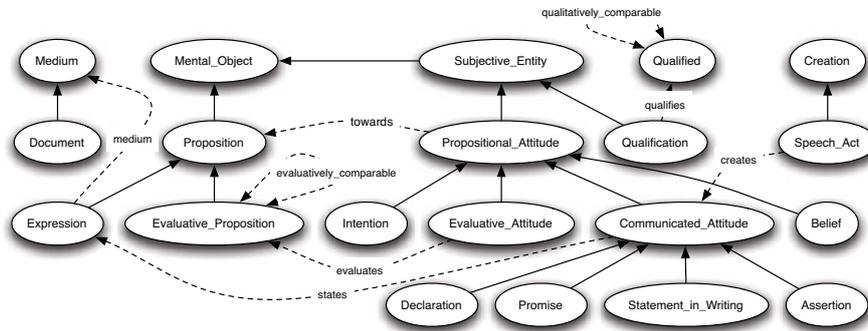


Figure 6.7: Propositions, Attitudes and Expressions.

used as a model for intelligent decision making and planning.<sup>24</sup> It is important to note that there is an *internalist* and an *externalist* way to use intentions and expectations. The external observer can only ascribe intentions and expectations to an agent based on his observed actions. The external observer will make assumptions about what is *normal*, or apply a *normative* standard for explaining the actions of the agent (Boer et al., 2005a; Boer, 2009).

The expression module covers a number of representational primitives necessary for dealing with Propositional\_Attitudes (Dahllöf, 1995). Many concepts and processes in legal reasoning and argumentation can only be explained in terms of propositional attitudes: a relational mental state connecting a person to a Proposition. However, in many applications of LKIF the attitude of the involved agents towards a proposition will not be relevant at all. For instance, fraud detection applications will only care to distinguish between potentially contradictory observations or expectations relating to the same propositional content. Examples of propositional attitudes are Belief, Intention, and Desire. Each is a component of a mental model, held by an Agent.

Communicated attitudes are held towards expressions: propositions which are externalised through some medium. Statement, Declaration, and Assertion are expressions communicated by one agent to one or more other agents. This classification is loosely based on Searle (1995). A prototypical example of a medium in a legal setting is e.g. the Document as a bearer of legally binding (normative) statements.

When propositions are used in reasoning they have an epistemic role, e.g. as Assumption, Cause, Expectation, Observation, Reason, Fact etc. The role a proposition plays within reasoning is dependent not only on the kind of reasoning, but also the level of trust as to the validity of the proposition, and the position in which it occurs (e.g. hypothesis vs. conclusion). In this respect, the expression module is intentionally left under-defined as a rigorous definition of epistemic roles would include their use in epistemological frameworks used in reasoning (see Section 5.5.2). As a consequence, the ontology allows one to ascribe an epistemic role to some entity, but cannot be used to infer it.

<sup>24</sup>Regardless of whether it is a psychologically plausible account of decision making. Dennett's notion of the *Intentional Stance* is interesting in this context (Dennett, 1987). Agents may do no more than occasionally apply the stance they adopt in assessing the actions of others to themselves.

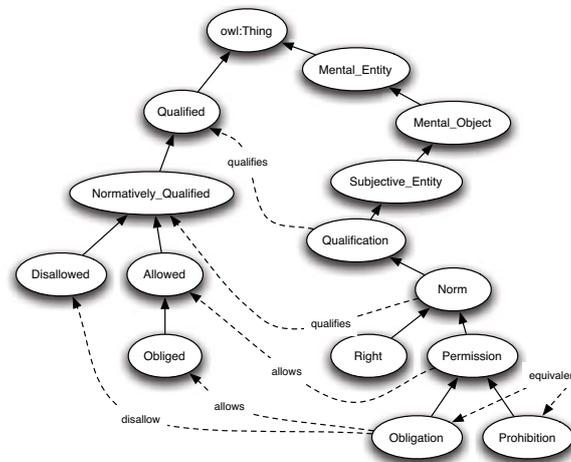


Figure 6.8: Qualifications and Norms

Evaluative\_Attributes express an evaluation of a proposition with respect to one or more other propositions, they express e.g. an evaluation, a value statement, value judgement, evaluative concept, etc. These attitudes express only the type of qualification which is an attitude towards the thing being evaluated, and not for instance the redness of a rose, as in Gangemi et al. (2002) and others. Of special interest is the Qualification, which is used to define norms as specified by Boer et al. (2005a). Analogous to the evaluative attitude, a qualification expresses a judgement. However, the subject of this judgement need not be a proposition, but can be any complex description (e.g. a situation).

### 6.3.3 The Legal Level

Legally relevant statements are created through public acts by both natural and legal persons. The legal status of the statement is dependent on both the kind of agent creating the statement, i.e. Natural\_Person vs. a Legislative\_Body, and the rights and powers attributed to the agent through mandates, assignments and delegations. At the legal level, the LKIF ontology introduces a comprehensive set of legal agents and actions, rights and powers (a modified version of Sartor (2006); Rubino et al. (2006)), typical legal roles, and concept definitions which allow us to express normative statements as defined in Boer et al. (2005a); Boer (2006); Boer et al. (2007a); Boer (2009).

The Norm is a *statement* that combines two meanings: it is *deontic*, in the sense that it is a qualification of the moral or legal acceptability of some thing, and it is *directive* in the sense that it commits the speaker to ensure that the addressee of the norm complies with it, e.g. through a sanction (Nuyts et al., 2005). These meanings do not have to occur together as it is perfectly possible to attach a moral qualification to something without directing anyone, and it is equally possible to issue a directive based on another reason than a moral or legal qualification (e.g. a warning).

The normatively qualified situations of LKIF are analogous to the generic

situations, or generic cases of Valente (1995). A norm applies to (or qualifies) a certain situation (the Qualified situation), allows a certain situation – the Obligated or Allowed situation – and disallows a certain situation – the Prohibited or Disallowed situation, see Figure 6.8. Allowed and disallowed situations are subclasses of the situation qualified by the norm. Furthermore, they by definition form a complete partition of this situation: a norm applies only to situations that are either mandated or prohibited.

As a consequence, norms of type Prohibition and Obligation are equivalent because they are simply two alternate ways to put the same thing into words: a prohibition to smoke is an obligation not to smoke. A Permission is different in that it allows a situation without prohibiting anything. In that case the logical complement of the allowed situation is some opposite qualified situation about which we only know that it is not obliged.

Where in other approaches, cf. CLO (Gangemi et al., 2005), a situation is the reification of some state of affairs, the normatively qualified situations in LKIF Core are *instantiated* by states of affairs: they are defined as class descriptions that represent a set of possible states of affairs. This means that a standard reasoner can infer whether some actual situation is subsumed under a generic situation, and thus whether norms exist that allows or disallows that situation. Similarly, a classifier will create an inferred hierarchy of situations, which enables a relatively straightforward resolution of *lex specialis* exceptions between norms (Boer et al., 2005a).

## 6.4 Discussion

This chapter presented a principled approach to ontology development in the legal domain, and described how it was applied in the development of the LKIF Core ontology. The main requirements for this knowledge representation ontology, as part of LKIF, are that it should be a resource for special, legal inference, and that it provides support for knowledge acquisition. The legal domain poses additional requirements on the ontology in that it presupposes a common sense grounding. Section 6.2.1 discusses several of the types of ontologies introduced in Section 5.4.1 and argues where they fall short in providing this grounding. Section 6.2.2 presents an approach based on insights from cognitive science and knowledge representation. This approach was incorporated in the middle-out methodology applied in the development of the ontology. The modules and main concepts of the LKIF Core ontology are discussed in Section 6.3. LKIF Core was represented using the OWL 2 DL language described in Chapter 3, and is available at the <http://purl.org/net/lkif-core> persistent URL.<sup>25</sup>

As LKIF Core was developed by a heterogeneous group of people, we specified a number of conventions to uphold during the representation of terms identified (Breuker et al., 2007). One of these is that classes should be represented using necessary & sufficient conditions as much as possible (i.e. by means of owl:equivalentClass statements). Using such ‘complete’ class definitions en-

<sup>25</sup>The home page for the ontology is currently hosted at <http://www.estrellaproject.org/lkif-core>. Separate modules of the ontology are reachable by adding the module name at the end of the persistent URL, e.g. <http://purl.org/net/lkif-core/action> for the action module.

sure the ability to infer the type of individuals. In retrospect, the extensive use of a combination of Generic Concept Inclusion axioms (equivalent class statements on existential restrictions) and inverse properties turned out to be quite taxing for (in our case) the Pellet reasoner. As at this time we used the reasoner primarily for debugging purposes, a single inconsistency in the TBox could cause the reasoner to stall, making it hard to debug the ontology. Although this problem was remedied by lifting some of the restrictions on classes in LKIF Core, it indicates that real time performance of reasoners on ontologies that use the full expressiveness of OWL 2 DL can still be improved.<sup>26</sup> The discussion of design patterns in the next chapter proposes a way to improve reusability by creating *summaries* of complex class definition patterns (see Section 7.3).

Using a large ontology such as LKIF Core in practice will inevitably be taxing in other ways as well (see Section 5.4). A reusing ontology will need to ascribe to its entire set of ontological commitments. Critical users may have trouble reconciling their own views with the ontology, where naive use may overlook important consequences of this commitment. Furthermore, the ontology involves a level of granularity and detail that may not directly fit a domain ontology. This can be overcome by importing available, compatible domain ontologies in a representation based on the LKIF ontology, provided that safe reuse is ascertained. A possibility is to create library of guaranteed compatible ontologies as part of the LKIF architecture. Reusing a core ontology requires a significant investment, even though improving reusability and facilitating knowledge acquisition are some of its principal purposes. For a knowledge representation ontology such as LKIF Core, the proof of the pudding is in the eating: the pay-off of adopting the ontology in reasoning should exceed the investment it requires.

The LKIF ontology has been successfully applied in the development of a generic architecture for legal reasoning (HARNES), as the main knowledge resource for performing normative assessment.<sup>27</sup> Hoekstra et al. (2007) evaluates the LKIF definition of Norm in the representation of part of an EU directive on traffic regulations. In van de Ven et al. (2008b,a) we adopted the definitions of the norms module to drive a Protégé plugin for normative assessment. This plugin has been tested to work with representations of both a toy domain of library regulations, and a part of the Hungarian tax law. It uses the Pellet reasoner to compute the subsumption hierarchy of generic cases qualified by the body of norms, and uses this information to generate a *lex specialis* exception hierarchy of norms.

It turned out that extending the ontology with concepts required for the domain of the regulations was relatively straightforward. However, specifying the contents of norms in terms of OWL constructs remained an arduous task because of the sheer complexity of the situations they govern. This experience suggests that the ontology should be extended with a module for expressing measurements, and led to the specification of design patterns that allow us to adequately capture *transactions* and other complex concepts (Hoekstra and Breuker (2008) and Section 7.2). Klarman et al. (2008) propose a way to combine the LKIF ontology with a versioning scheme that corresponds to a specification

<sup>26</sup>Thanks to Taowei David Wang for pointing this out, see <http://lists.owlidl.com/pipermail/pellet-users/2007-February/001263.html>

<sup>27</sup>HARNES: Hybrid Architecture for Reasoning with Norms by Employing Semantic Web Standards

---

of validity intervals for legal sources.

The following chapter gives an overview of several design patterns identified in the construction of the LKIF ontology. The purpose of this overview is twofold. First, and foremost, it illustrates the wide applicability of design patterns across domains, and secondly, it provides insight in the definition of some of the central concepts in the LKIF ontology.