LKIF Core : Principled Ontology Development for the Legal Domain

Rinke HOEKSTRA [a], Joost BREUKER [a], Marcello DI BELLO [b], and Alexander BOER [a]

[a] *Leibniz Center for Law, University of Amsterdam*
breuker@science.uva.nl, hoekstra@uva.nl, aboer@uva.nl
[b] *Department of Philosophy, Stanford University*
mdibello@stanford.edu

Abstract. In this paper we describe a legal core ontology that is part of the *Legal Knowledge Interchange Format*: a knowledge representation formalism that enables the translation of legal knowledge bases written in different representation formats and formalisms. A legal (core) ontology can play an important role in the translation of existing legal knowledge bases to other representation formats, in particular as the basis for articulate knowledge serving. This requires that the ontology has a firm grounding in commonsense and is developed in a principled manner. We describe the theory and methodology underlying the LKIF core ontology, compare it with other ontologies, introduce the concepts it defines, and discuss its use in the formalisation of an EU directive.

Keywords. ontology, legal ontology, ontological engineering, legal concept, LKIF, knowledge representation, framework, methodology, common sense

Introduction

In this paper we describe a legal core ontology that is part of the *Legal Knowledge Interchange Format*, currently being developed within the Estrella project.[1] This interchange format has two main roles: it enables the translation between legal knowledge bases written in different representation formats and formalisms and secondly, it is a knowledge representation formalism in its own right, as part of a larger architecture for developing legal knowledge systems. These use-cases expose LKIF to the classical trade-off between tractability and expressiveness, as in e.g. KIF (Knowledge Interchange Format, [25]). An additional requirement is that LKIF should comply with current Semantic Web standards to enable legal information serving via the web: LKIF builds on a combination of OWL DL and a rule formalism, e.g. RIF-BLD[2], offering a classical hybrid solution. How

[1] Estrella is a 6th European Framework project (IST-2004-027665). The views and work reported here are those of the authors. See `http://www.estrellaproject.org`

[2] The Rule Interchange Format-Basic Logic Dialect (see `http://www.w3.org/2005/rules/wg/bld/draft-2007-10-30`). In the first version of LKIF [7] it was proposed that its most basic layer should consist of OWL DL and SWRL , complying with standards of the Semantic Web. However, development on SWRL has been replaced by a proposal for a rule interchange format, RIF that is intended to become the

these two formalisms are to be combined still is an important issue in the specification of LKIF.

However, OWL DL and RIF do not make LKIF a formalism tuned to *legal* knowledge and reasoning: how do we get the 'L' into LKIF? "Legalising" LKIF can be done in two ways: either extend LKIF by specifying typical legal *tasks*, or by including legal *knowledge*. Tasks, or more precisely, problem solving methods (PSM), can be specified in OWL DL and rules as a *meta-component* that controls reasoning. This is exemplified by the classical approach of constructing knowledge system 'shells' [39,16], or a library containing formally specified PSM [9]. Typical legal tasks are legal assessment – e.g. in judging cases, or in legal advice (legal question answering) – and legal planning [58].

However, the PSMs that are suitable for executing these tasks are not necessarily different from those used for the same tasks in other domains. In other words, although these reasoning architectures may be very useful for legal domains, they do not really 'legalise' LKIF. For example, considering practice in legal courts, one could argue that *argumentation* is the most typical and exclusive legal task, even if argumentation is only a small part of what goes on in law and legal issues. However, argumentation by itself is not typical for law at all: it plays a role in *any* goal directed reasoning, i.e. problem solving.

Solving problems is done by means of constructing and selecting potential solutions (hypotheses), and testing these against evidence or requirements. Support and proofs of consistency and compliance are in fact arguments that reject or maintain the hypotheses: arguments are part of any PSM. Nonetheless, a significant difference from argumentation in general is that in law, and particularly in court, argumentation is achieved through dialogues. Moreover, these dialogues are not co-operative but *adversarial*: they are disputes. The Carneades model of legal argumentation, developed by [26], is implemented as part of LKIF in Estrella. Carneades combines the use of argumentation schemes with a highly expressive rules language, LKIF Rules, as specified in [7].[3]

In this article we are concerned with the second way of legalising LKIF, i.e. by capturing the *content* of legal reasoning: legal knowledge. In every domain, reasoning can only be performed based on some understanding of that domain. The primary source for this understanding lies in terminological knowledge, which enables us, but also machines, to recognise occurrences such as entities or events as belonging to a particular category (concepts, classes). In other words, terminological knowledge enables an agent to give 'meaning' to these occurrences.

In knowledge bases, terminological knowledge is usually expressed as an *ontology*, and description logic based formalisms such as OWL are specifically designed for this purpose. As LKIF is intended to cover all legal domains, an LKIF ontology should focus on concepts that are typical for the whole range of legal domains, i.e. they should be sufficiently general and abstract. This kind of ontology is called a *core* ontology. A core ontology mediates between general terms that cover all domains – a *top* or upper ontology

---

common denominator of several well defined dialects. The underlying intentions of this proposal fit well with those of LKIF , whose interchange function can be supported through alignment with this initiative. However, a disadvantage is that it will not only take some time before this proposal becomes a standard, but also that RIF compliant technology is most likely to cater for specific dialects which are not necessarily suitable as basis for LKIF. The Estrella project foresees a revision and refinement cycle in 2008 in which these new developments will be considered.

[3] The LKIF Rules formalism is significantly more expressive than RIF. We are currently investigating the possibilities to identify a first-order subset of LKIF Rules which can be mapped, and translated to RIF.

– and domain ontologies. LKIF, a combination of knowledge representation formalisms, is filled and extended by such a core ontology: LKIF Core.

In the context of LKIF, the ontology provides support in several ways. 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*, and lastly, a terminological framework can facilitate the *exchange* of knowledge across multiple knowledge bases.

Resource for special, legal inference. Typical, 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. These implied inferences are relevant when reasoning with norms and cases.

It is important to distinguish between the automatic elicitation of implicit, or implied knowledge and the control over which knowledge is used when, where and how. For the former we reserve the term 'inference', for the latter we use 'reasoning'. This distinction coincides with the distinction between sanctioned inferences and recommended inference, cf. [18]. 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 using the definitions in an ontology and the (general) inference engine that goes with the representation formalism[4] to produce specialised inference is e.g. temporal reasoning based on Allen's ontology of time, cf. [1]. To enable special inference, terms should be highly interrelated and form a coherent 'cluster', with no or little external dependencies ([33], and Section 2). A good example of such a cluster for legal reasoning is formed by the terms that denote deontic qualifications. These clusters 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. 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.

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 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.

[4] For an ontology cast in OWL DL these inference engines are description classifiers, e.g. Pellet, `http://pellet.owldl.com/`

The following sections describe the theoretical and methodological framework against which the LKIF core ontology has been developed. Section 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 2. Section 3 extends this methodology by introducing a distinction between ontologies and *frameworks*, followed by a discussion of the modules and most important concepts of LKIF Core in Section 4. Section 5 gives an example of how the ontology can be used in the formalisation of a regulation.

1. Points of View: Other Ontologies

It is not recommended to develop any ontology from scratch, and this holds especially for top and core ontologies. The topmost layers of such an ontology will inevitably need to contain the highly abstract categories underlying our knowledge: i.e. terms that have been studied for centuries by philosophers [54]. Therefore, it seems advisable to consult publicly available top-ontologies for re-usable definitions (see also Section 2). In this section, we show that even at this highly abstract level, the point of view taken in the development of an ontology has a large impact on *how*, and *which* categories are defined.

LKIF Core is intended to contain a core set of definitions for modelling legal terms. Law can be viewed as an instrument used by the legal and political system to identify and control situations and events in social interaction. By far the bulk of social situations – in family life, work, transport, property, crime, etc. – does not need to be 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. However, this terminology needs to be grounded in the kinds of events and states to be handled by law: even this legal technical terminology cannot be disconnected from day to day life. [5] In other words, the LKIF ontology should have a common sense grounding.

An additional requirement for LKIF is that it should be based on Semantic Web technology. The primary target –or Tim Berners-Lee's dream – of the Semantic Web is to achieve mutual understanding between web services and human users. This requires at least some common sense understanding by the machine, and a common sense oriented top-ontology would be a first step. In short, the common sense perspective is also applicable to any serious endeavour towards a Semantic Web.

Given this perspective, we expected that (at least parts of) existing ontologies would be re-usable, ranging from a source of inspiration to straightforward import of definitions. This would hold, in particular, for top ontologies that include legal terms, as for instance listed in [15]. Unfortunately, it turned out that the amount of re-use and inspiration was rather limited. We consulted several ontologies and evaluated them for their potential contribution for creating a coherent top for LKIF Core, and specifically for common sense and legal terms already represented. Additional requirements were their suitability

[5] One can object that this translation is catered for by legal professionals. However, this role is limited to special occasions – courts – or domains – e.g. tax. The default assumption that every citizen knows the law still holds: thus, the majority of legislation and legal practice should be understandable by the public.

for enabling the primary roles of the LKIF ontology: knowledge acquisition, and special inference.

1.1. Suggested Upper Merged Ontology

The SUMO[6] ontology [47] was input to a large scale effort to build a standard upper ontology (SUO) and brings together insights from engineering, philosophy – e.g. Sowa's upper ontology [54] – and information science. SUMO provides definitions for general purpose terms, and can be used as a unifying framework for more specific domain level ontologies. It was created by merging existing ontological content of a wide variety into a coherent whole.

As a starting point for the LKIF core ontology, SUMO has several drawbacks. First of all, it is specifically intended to be an upper ontology and thus does not readily provide definitions of terms relevant to the legal field – e.g. mental and social entities are poorly represented. Furthermore, because of the way in which SUMO is constructed, it has a strong bias towards the more abstract, theoretical insights coming from engineering and philosophy: it does not have a common-sense basis. Lastly, SUMO is specified using KIF, and although a translation to OWL Full is available, it would take considerable effort to turn it into a manageable OWL DL ontology.

1.2. Descriptive Ontology for Linguistic and Cognitive Engineering

DOLCE[7] was developed as part of the WonderWeb library of foundational ontologies, cf. [40,22]. This ontology was developed 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. 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. Like SUMO, DOLCE was originally specified in the highly expressive languages FOL and KIF, and its OWL DL representation (DOLCE-Lite) is more restrictive, e.g. it does not consider modality, temporal indexing and relation composition.

DOLCE is very much an ontology in the traditional philosophical sense: Ontology as the philosophical study of existence. We argue that (as is indeed pointed out by various authors, e.g. [30]), the aim of Ontology as studied by philosophers differs significantly from ontologies as developed in AI and the semantic web. Compared to the philosophical perspective, we prefer a broader interpretation of what an ontology is: ontologies should reflect our *knowledge*, i.e. the categories, things, concepts with which we are able to understand the world. This interpretation reflects a knowledge engineering perspective in which an ontology is a description of the terms used in a knowledge base, and not of reality. In fact, many ontologies in AI do not aim to be ontologically sound in the philosophical sense. The DOLCE approach differs from this perspective in two significant ways. Firstly, its perspective is philosophical with respect to its *content*, i.e. aimed at re-

[6] http://ontology.teknowledge.com, submitted to the IEEE Standard Upper Ontology Working Group, http://suo.ieee.org/

[7] http://www.loa-cnr.it/DOLCE.html

ality. And secondly, as a consequence of the philosophical perspective, it is rather an extension of the knowledge representation formalism by introducing additional categories (viz. OWL's classes and properties) at the *ontological level* [32], than a model expressed using that formalism. In other words, the meta-level character of DOLCE means that the ontology is not a representation of knowledge itself, but only of the terms used for describing that knowledge. It is because of this discrepancy between the two views that it is sometimes even argued that the Web Ontology Language (OWL, [3]) is not suitable for capturing philosophical ontology, cf. [4].[8]

The DOLCE ontology is *descriptive*, and is based on the stance that "the *surface structure* of natural language and human cognition"[9] 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 explicit commonsense perspective, compared to the science perspective of SUMO. However, the claim that this surface structure has any bearing on commonsense is not based on evidence. The methodological commitment to the *surface* structure of language and cognition has, in our view, lead to an ontology that provides an intricate framework of theoretical notions, rather than an ontology that directly *captures* commonsense.

1.2.1. Core Legal Ontology

Over the years, DOLCE has been extended in several ways, the most relevant being DOLCE+ and CLO. DOLCE+ is an extension of DOLCE with a theory on descriptions and situations (also called D&S, [23]). CLO, the Core Legal Ontology [24], organises legal concepts and relations based on the formal properties defined in DOLCE+. It was designed to support both the definition of domain ontologies and a juridical Wordnet, and the design of legal decision support system. To a large extent these goals are in accordance 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 *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.

The legal system as description, or rather *pre*scription, on reality is not new, cf. [58, others]. However, the CLO distinction between descriptions and situations is rather one dimensional and does not commit to an ontological view of the *kinds* of descriptions involved. It confounds the distinction between representation and the *represented* with rep-

[8]As we discussed before OWL DL is particularly suited for classification and thus has an ontological bias with respect to this task: its primitives are classes and properties. Since we commit to the stance that defining what things *are* essentially boils down to classification, this bias is entirely instrumental to our needs.

[9]Emphasis by the authors, [40]

resentation and *reality* and thus, although it allows for multiple levels of reification, does not provide ontological categories for these levels. In fact, according to the knowledge engineering perspective outlined in our discussion of DOLCE, anything in an ontology is a representation of something in reality.

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.

1.3. CYC

CYC is a huge knowledge base of commonsense concepts [38]. 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" [42, p.2].

The idea is that, when enough commonsense knowledge is represented, and a certain threshold is reached, a quantum-leap ("The Singularity")[10] 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 common sense knowledge [44,45]. 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.[11]

The upper part of the CYC ontology is thus claimed to have a common-sense 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. 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 knowledge engineer simply put his own commonsense knowledge, i.e. his understanding of the world, into the CYCL formalism. In other words, CYC captures common knowledge, and not necessarily commonsense itself.

1.4. LRI-Core

Thus far, the ontologies we reviewed do not appear to meet our requirements for the top structure of a legal core ontology. Concepts that are typical for law are either scarce and under specified, or overly theoretical, which limits the possibilities for reuse. Although in the past few years the ontologies underwent significant changes and extensions, this

[10] Indeed, CYC is a much-hyped project, and has received a lot of criticism because of it.

[11] Although the online game FACTory, for teaching CYC new things has been set-up by Cycorp, it does not seem to be online at the moment.

result is in line with the outcome of a similar review about three years ago ([11]). In particular, the requirement that a legal core ontology should provide the grounding of a common sense point of view was and still is absent. Where a common sense perspective is claimed, it is not motivated, explained or substantiated. This earlier review motivated the 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. [12] This ontology, *LRI-Core*, had largely the same requirements as LKIF Core, cf. [12,11]. What sets LRI-Core apart from other ontology efforts is that it is based on empirical studies and insights in cognitive science.

The search for the roots of common sense starts with an evolutionary reconstruction. Organisms have developed the abilities to perceive and move in order to deal with a changing and potentially dangerous world. As a consequence, rudimentary processing of sensory data into coherent interpretations of the environment evolved into more sophisticated neural wiring. This neural wiring is to a large extent genetically encoded. For instance, in the occipital lobe of mammals and birds – 2 neurons away from the retina – straight lines and angles are identified by combining sensory data from adjacent rods and cones in the retina: these sensors themselves do not form straight lines at all. In more complex brains, sensory data are first abstracted to basic properties, after which they are coherently synthesised and put into a spatial and historical context. This indicates that anatomical structures are the primary data processors: a hard wired knowledge representation. More abstract and context sensitive processing – synthesising – is genetically inherited as instincts; i.e. it requires some combination of constraining dispositions and experiences to 'program' a relatively flexible brain. In other words, the typical competencies of our genetic ancestors give insight in what the roots of our common sense are. And, these roots may well be hidden too deep to be reachable through introspection. However, some global and primary distinctions can be made on the basis of well established facts and strong beliefs in cognitive science.

The story starts as follows. From within a spatial and temporal embodiment, our physical environment (earth) is a very stable place – the notable exceptions being day-and-night cycles, weather and geological perturbations. The advent of organisms with the ability to move around, has created a prominent role for the perception of organism induced change: perceiving other organisms is highly relevant. Firstly, they may present an opportunity for *reproduction* and *metabolism* (eating); the main processes sustaining any organism. Secondly, they may present a direct threat to an organisms' existence. In both of these cases it pays to single out change from its relatively static background using *attentional mechanisms* which are extremely alert to change, and its interpretation.

The result is a prominent distinction in cognition between 'background' and 'foreground'. The background is maintained by two major elements: spatial arrangements of objects and historical continuity. Historical continuity is the requirement for a(n episodic) memory: the ability to perceive stability. Spatial stability is maintained by the fact that physical objects keep their position, unless subjected to the exertion of some force. [13] Changes – events – occur against this canvas of temporal and spatial positions. Time is important as it is the speed of a change that determines whether it becomes foreground.

---

[12] Electronic Court: Judicial IT-based Management. e-Court was a European 5th Framework project, IST 2000-28199.

[13] That stable positions are ensured for (heavy) objects is not only due to their inertia (mass), but probably more to counteracting forces (gravity, friction).

What implications did this story have for the construction of LRI-Core?

- Our knowledge serves to interpret occurrences in the world. This is reflected by the distinction between concepts on one hand, and individuals and their occurrence (instance) on the other hand. [14]
- Space and time serve as the most stable reference framework, i.e. space and time are assumed to be unaffected by anything that happens in the world. [15]
- The primary distinction between static and dynamic elements in the (physical) world is reflected by the distinction between objects and processes. Processes give a causal explanation of changes.
- Objects have extensions in space, i.e. they 'consume' some space; processes have extensions in time, i.e. they consume time. Processes are contained by objects.
- Space and time are viewed both as properties of objects (i.e. their extension) and as positions, defined by spatial referents. Position is not an inherent property of objects, i.e. changing the position of an object does not change the object itself.

Thus far we have only outlined the implications of the most elementary distinctions that a sensitive organism needs to cope with in a *physical* world, or rather: in a physically interpreted world. Very recent, in evolutionary terms, animals (mammals) have developed the ability to attribute a mind to their fellow animals, i.e. they see their behaviour as intentional. Changes are no longer only due to physical causation – i.e. they simply happen given certain circumstances – but can be attributed to an intentional agent as well. This requires a distinction between *process* and *action*. With each action an agent intends a change, i.e. he initiates processes that bring about the intended change.

Social behaviour is exhibited by many 'lower' animals such as ants and bees. However, this behaviour is too 'mechanical' to attribute it to an intentional stance [19]. The typical proof for the presence of an ability to take an intentional stance is the capacity of *anti*-social behaviour: e.g. fooling a fellow animal. In fact, many species of monkeys are capable of this behaviour. Paradoxically enough, taking into account the mental state of other animals has preceded the ability to model one's own mind. Being able to consciously reflect on your own mind was long thought to be a skill exclusive to humans. However, it now appears that our closest contemporary species – chimpanzees, bonobo's – have self awareness as well.

Nonetheless, interpreting the mental world is a recent accomplishment, and we hold the position that mental models use the conceptualisations of the physical world we have acquired earlier in evolution. We talk about mental *processes* [16] (e.g. forgetting) and mental *objects* (e.g. thoughts) in the same terms as we we talk about the physical world. In other words, our mental world is a metaphor of the physical world. It may well be that we even attained a world of total abstractions (mathematics) through metaphors, as is argued by Lakoff and Núñez, [36].

Social awareness and self awareness are the roots for complex social behaviour. They allow us to enhance the predictability of our environment through planning and control. Repeated execution of co-operative plans can render them well established and

---

[14] This is more or less already reflected in OWL's basic categories. OWL does not distinguish between an individual and its occurrence (instance). LRI-Core has a basic terminology to distinguish occurrences.

[15] Gravity is also stable, but it coincides with a spatial dimensions, Moreover, we cannot perceive distances in gravity; distances are the basis for identifying positions.

[16] Also: mental actions, e.g. in (trying) to control one's thoughts.

even ‘institutionalised’. These fixed co-operative plans make extensive use of *roles* to specify expected behaviour. Moving from plans to roles, we arrive at social organisation. The ability to play a certain role, expresses a (recognised) competence, sometimes acknowledged as a social ‘position’. In LRI-Core, roles are the mental objects by which we create social structures.

In summary, LRI-Core distinguishes five ‘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: roles
4. An *abstract world*, which contains only formal objects.

Although LRI-Core was intended as a core ontology for law, the number of typical legal concepts in this ontology has remained very limited. This is due to the fact that most of the effort in its specification was directed towards higher level categories. On the other hand, its methodological grounding in cognitive science, and explicit common sense perspective makes it a more likely candidate for fulfilling the requirements for the LKIF ontology than the other ontologies we discussed so far.

LRI-Core can in many ways be seen as the direct precursor of the LKIF ontology.[17] However, LKIF Core is not simply a specialisation of LRI-Core, but rather a massive overhaul. Although the perspective and main distinctions of this ontology were used as inspiration, the LKIF ontology was built from the ground up. In the following section we reformulate the cognitive science perspective as basis for the methodology we adopted in the construction of the ontology.

2. Methodology

The construction of LKIF Core is guided by a combination of methodologies for ontology engineering. Already in the mid-nineties, the need for a well-founded ontology development methodology was recognised, most notably by [28,29,57,56] and later [21]. These methodologies follow in the footsteps of earlier experiences in knowledge acquisition, such as the CommonKADS approach [52] and others, but also considerations from *naive physics* and *cognitive science*, such as [33] and [35], respectively.

Hayes, in [33], describes an approach to the development of a large-scale knowledge base of naive physics. Instead of rather metaphysical top-down construction, his approach starts with the identification of relatively independent *clusters* of closely related concepts. These clusters can be integrated at a later stage, or used in varying combinations allowing for greater flexibility than monolithic ontologies. Furthermore, by constraining (initial) development to clusters, the various – often competing – requirements for the ontology are easier to manage.

Whereas the domain of [33]’s proposal concerns the relatively well-structured domain of physics, the combination of commonsense and law does not readily provide an

[17] This is no really surprising as there exists a large personal overlap between the developers of LRI-Core and LKIF Core.

obvious starting point for the identification of clusters. In other words, for LKIF Core, we cannot carve-up clusters from a pre-established middle ground of commonsense and legal terms. Furthermore, the field does not provide a relatively stable top level from which top-down development could originate.

In [57], who are the first to use the term ‘middle-out’ in the context of ontology development, it is stressed that the most ‘basic’ terms in each cluster should be defined before moving on to more abstract and more specific terms within a cluster. The notion of this basic level is taken from Lakoff ([35]), who describes a theory of categorisation in human cognition. Most relevant within the context of ontology engineering [57,35, p. 12 and 13] are *basic-level categorisation*, *basic-level primacy* and *functional embodiment*. Categories are organised so that the categories that are cognitively basic are ‘in the middle’ of a taxonomy, generalisation proceeds ‘upwards’ from this basic level and specialisation proceeds ‘downwards’. Furthermore, these categories are functionally and epistemologically primary with respect to (amongst others) knowledge organisation, ease of cognitive processing and ease of linguistic expression. Basic level concepts are used automatically, unconsciously, and without noticeable effort as part of normal functioning. They have a different, and more important psychological status than those that are only thought about consciously. This approach thus fits well with the theoretical considerations of the previous section.

For the purpose of the LKIF ontology, we have made slight adjustments to the methodology of [33,56]. We established design criteria for the development of the LKIF ontology based on [27,56]. These criteria were implemented throughout the following phases: identify *purpose and scope*, ontology *capture* and *coding*, *integration* with existing ontologies and *evaluation*.

2.1. Ontology Capture

The LKIF Core ontology should contain ‘basic concepts of law’. However, it depends on the (potential) users what kind of vocabulary is aimed at, and which concepts are *basic*. 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 exists a sufficient common understanding to treat them more or less as similar, cf. [37]. Nonetheless, a number of basic terms have a specific legal-technical meaning, such as ‘liability’ and ‘legal fact’. Technical terms were included because they capture the ‘essential’, abstract meaning of terms in law. Furthermore, these terms can be used to structure the relations between more generally understood legal terms.

The Estrella consortium includes representatives of these three kinds of experts. Each partner was asked to supply their ‘top-20’ of legal concepts. Combined with terms we collected from literature (jurisprudence and legal text-books) we obtained a list of about 250 terms. As such a number is unmanageable as a basic set for modelling, we asked partners to assess each term from this list on five scales: level of *abstraction*, *relevance* for the legal domain, the degree to which a term is *legal* rather than *commonsense*, 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 those re-used from other ontologies (see section 1), and formed the

Table 1. Ten highest scoring terms for *importance*, *abstractness*, and *legal relevance*

Importance	Abstractness	Legal Relevance
Law	Deontic operator	Civil law
Right	Law	Law
Jurisdiction	Norm	Legal consequence
Permission	Obligative Right	Legislation
Prohibition	Permissive Right	Obligation
Rule	Power	Right
Sanction	Right	Authority
Violation	Rule	Deontic operator
Power	Time	Duty
Duty	Anancastic Rule	Jurisdiction
Legal Position	Existential Initiation	Legal Fact
Norm	Existential Termination	Legal Person
Obligation	Potestative Right	Legal Position
Permissive Right	Productive Char.	Legal Procedure
Argument	Absolute Obl. Right	Liability

basis for the identification of clusters and the development of the LKIF Core ontology. Table 1 shows the ten highest scoring terms for three of the scales.

3. Frameworks and Ontologies

Many of the terms we gathered are not necessarily suited for inclusion in the core ontology. Even if scores on the five scales indicate that a term is of the utmost importance, this does not immediately mean that a term, and the relations it has, is of *ontological* relevance. We adhere to a rather restrictive view on what an ontology should contain: terminological knowledge, i.e. intensional definitions of concepts, represented as classes with which we interpret the world. Any definition that does not meet this rule should not be part of an ontology. For this reason, we have advocated a distinction between ontologies and so-called *frameworks* in earlier papers, i.e. [13,34].

The distinction between terminological knowledge (T-Box) and assertional knowledge (A-Box) has already been around for a long time. As a rule, terminological knowledge is generic knowledge while assertional knowledge describes the (actual) state of some world: situations and events. However, these asserted states can become generalised into typical patterns related to particular situations. To be sure, if experiences reoccur and have a justifiable structure, it might evidently pay to store these structures as generic descriptions, because they capture a predictable course of events. Eating in a restaurant is a typical example, and served in the Seventies to illustrate the notion of knowledge represented by scripts [51] or 'frames' [43]. For representations of this kind of generic knowledge, which is indeed rooted in terminological knowledge, we use the term *framework*, as these representations are structurally different from ontologies.

Ontologies capture what things *are* in and of themselves, whereas frameworks capture the systematic co-occurrence, the structural *relations*, something has with other things. Where ontologies consist of definitions of terms that have 'natural' subsumption ('is-a') relations, frameworks describe how activities are causally or intentionally related, or how objects are spatially and functionally configured.

Arguably, at a formal level (e.g. in OWL) these are indistinguishable: every class in OWL is defined in relation to other classes, it cannot be otherwise. This distinction therefore is conceptual: it does not map easily onto representation formalisms. The conceptual distinction more or less reflects the difference between intrinsic and accidental relations. We distinguish between those relations that make what a concept *is* and relations that place a concept in a particular frame of reference, or *context*. When a context is recurring and sufficiently stereotypical, it might well pay to represent its characteristic features. These features are not real 'properties', as they do not 'define' a concept, but merely enable its recognition.

Take for example a *hammer*, its composition of head and shaft is not by accident: it is this particular combination that allows the hammer to be used 'as a hammer'. However, the mereological relation between the hamer and its composites is not part of its ontological definition. Many different kinds of hammers exist, eg. sledgehammers, mallets, conveniently shaped stones etc., each of which differs ever so slightly in the nature of its composites and the relations between them. But they are all hammers. It is therefore only the function of the hammer as an instrument that defines what a hammer *is*, its mereological properties are merely circumstantial evidence.

From a methodological point of view, this allows us to introduce a rule of thumb: a combination of concepts and relations, as in e.g. a class restriction, is only to be considered part of an ontology when this particular combination is systematic, independent of context. A possible second consideration is inspired by the limitations imposed on admissible structures by the knowledge representation formalism. Of course this is a rather practical consideration, but nonetheless useful in this OWL-ified world. If the primary task of ontology is to describe what things are, then a representation formalism specifically tailored to classification can be a fair benchmark for determining whether the kind of knowledge you want to express should be considered part of ontology or not. Given the limited expressiveness of OWL , many frameworks are more easily represented using a rule formalism as they often require the use of variables. Frameworks often imply epistemic roles which require reasoning architectures that go beyond the services provided by OWL DL reasoners (e.g. when they require meta-level reasoning). This could indicate a correlation between the conceptual distinction and representation formalisms. However, the two perspectives should not be confounded. Frameworks belong to the T-Box of a knowledge representation system, independent of whether the T-Box is based on a DL formalism or not.

We have distinguished the following types of frameworks:

3.1. Situational frameworks

Situational frameworks are stereotypical structures of plans for achieving some goal in a recurrent context, e.g. making coffee. However, these plans can involve transactions in which more than one actor participates. For instance, the definition of `Eating-in-a-restaurant`[18] takes the dependencies between actions of clients (ordering, paying) and service personnel (noting, serving) as its major structure. This is the internal structure of the concept, but it usually does not make sense to create class-subclass relations between such frame-like concepts. The `Eating-in-a-restaurant`

[18] In the following all concepts will start with a capital letter, properties and relations will not

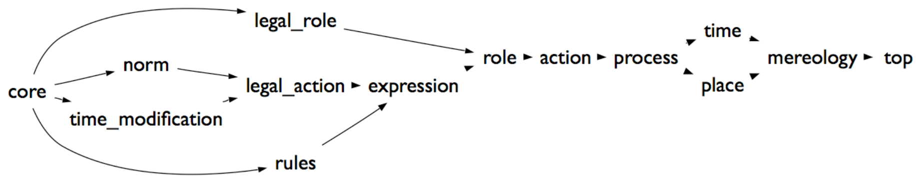

Figure 1. Dependencies between LKIFCore modules.

is not some *natural* sub-class of `Eating`. It refers to some typical model of how eating is put in the context of a restaurant. We can introduce a proliferation of all contexts of eating, such as `Eating-at-home`, `Eating-with-family`, etc. but these contexts do not fundamentally differ, cf. [5,11]. In the legal world, such situational frameworks may be pre-scribed in articles of procedural ('formal') law. Although *stereotypical* plans ('customs') and *prescribed* plans may differ in their justification – rationality vs. authority – their representation is largely analogous. Similarly, legal norms combine generic situation descriptions with some specific state or action. The description is qualified by a deontic term. For instance, the norm that "vehicles should keep to the right of the road" states that the situation in which a vehicle keeps to the right is obliged.

3.2. Mereological frameworks

Many entities, both objects and processes often have parts: they are *composites*. It is tempting to include a mereological (part-of) view in the definition of a concept. For instance, defining a car as having at least three, and usually four wheels, and at least one motor. However, a full *structural* description of all its parts and connections goes beyond what a car *essentially* is. Mereological frameworks appear under a large diversity of names: structural models, configurations, designs, etc. Arguably, the distinction between a mereological framework and a defining description of a term (ontology) is sometimes very thin. For instance, if we want to describe a bicycle as distinct from a tricycle, it is necessary to use the cardinality of the wheels as defining properties as these are *central* to the nature of the bicycle. On the other hand, the number of branches a tree might have hardly provides any information as to what a tree *is*.

3.3. Epistemological frameworks

Inference structures are often represented as epistemological frameworks of interdependencies between reasoning steps. Typical examples are the problem solving methods (PSM) found in libraries of problem solving components [9,46,52]. A problem solving method is not only a break-down of a problem, but also provides control over the making of inferences by assessing success and failure in arriving at the (sub)goals. PSMs have two major components: some method for selecting or generating potential solutions (hypotheses), and some methods for testing whether the solutions hold. Whether they hold may be due to the fact that they satisfy all the specified requirements (constraints) or whether they correspond with ('explain') empirical data.

This focus on the *use* of knowledge, its epistemological *status* (e.g. hypothesis vs. conclusion) and the dependencies between distinct steps in a methodology is characteristic for epistemological frameworks. Epistemological frameworks can be more abstract

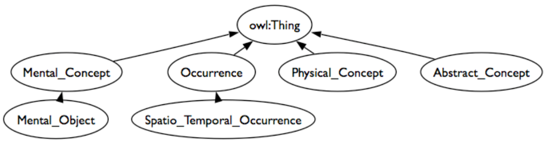

Figure 2. Concepts defined in the Top module.

than PSMs. For instance, the Functional Ontology of Law, which is presented as a core ontology, is an epistemological framework that describes the role of law as a control system in society [58,14].

4. Ontology Modules

The preceding sections introduce our requirements for LKIF Core and specify a principled approach to (legal) ontology development, which is based on insights from cognitive science and uses a well established methodology. With these considerations in hand, the LKIF ontology was initially designed as a collection of eight ontology modules: *expression*, *norm*, *process*, *action*, *role*, *place*, *time* and *mereology*, cf. [10]. This collection was later extended with a top ontology, two more ontology modules (*legal_action*, *legal_role*), and two frameworks (*time_modification* and *rules*), see Figure 1, [34,13]. Each of these modules contains a relatively independent cluster of concepts, represented using OWL DL in a middle-out fashion: for each cluster the most central concepts were represented first.[19]

We can distinguish three layers in the ontology: the *top* level (Section 4.1), the *intentional* level (Section 4.2) and the *legal* level (Section 4.3). These layers correspond to the different 'stances' one can adopt towards a domain of discourse, and are inspired by the work of Dennett ([19]). He identified three stances we can adopt for explaining phenomena in the world: the *physical* stance, used for explaining behaviour in terms of 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. The LKIF ontology thus adds a *legal* layer, containing concepts that are only sensible from a legal perspective. At each level, concepts are expressed in terms of concepts defined at a lower level, adding new organising structures (such as properties) where necessary. This methodology ensures the modular set-up of the ontology, as it allows users to commit to alternative theories at any level.

4.1. 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, part-hood and change are indispensable in a description of even the simplest legal account. The top level clusters of the

[19] We used both TopBraid Composer (`http://www.topbraidcomposer.com`) and Protégé 3.2/4.0a (`http://protege.stanford.edu`).

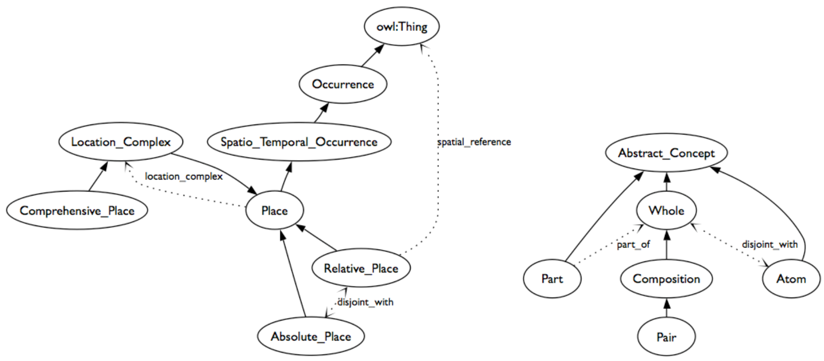

Figure 3. Place and Mereology related concepts.

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 concepts, and occurrences (Figure 2).

Mereological relations allow us to define parts and wholes, allow for expressing a systems-oriented view on concepts, such as functional decompositions, and containment (Figure 3). Furthermore, they form the basis for definitions of *places* (location) and moments and intervals in *time*.

The ontology for places in LKIF Core is based on the work of [20], and adopts a distinction between *relative* places and *absolute* places, which goes back to Newton. Whereas 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. See figure 3 for an overview of concepts defined in the place module. A `Location_Complex` is a set of places that share a reference location.

Of the properties defined in this module, `meet` is the most basic as it is used to define many of the other properties such as abut, cover, coincide etc. See [13,20] for a more in depth discussion of these and other relations. The current version of the ontology of places does not define concepts and relations that can be used to express direction and orientation.

Closely related to the theory of places of [20] is Allen's theory of time [1,2]. We adopt his theory, 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 4).

The relations between temporal occurrences are what defines time. Like [20], [1] adopts the `meet` relation to define two immediately adjacent temporal occurrences. We call this relation `immediately_before`, as the temporal meet relation holds only in one direction, and is asymmetric. The property is used to define other temporal relations such as `before`, `after`, `during`, etc.

With these classes and properties in hand, we introduce concepts of (involuntary) change. The process ontology relies on descriptions of time and place for the representation of duration and location of changes. A `Change` is essentially a difference between the situation before and after the change. It can be a functionally coherent aggregate

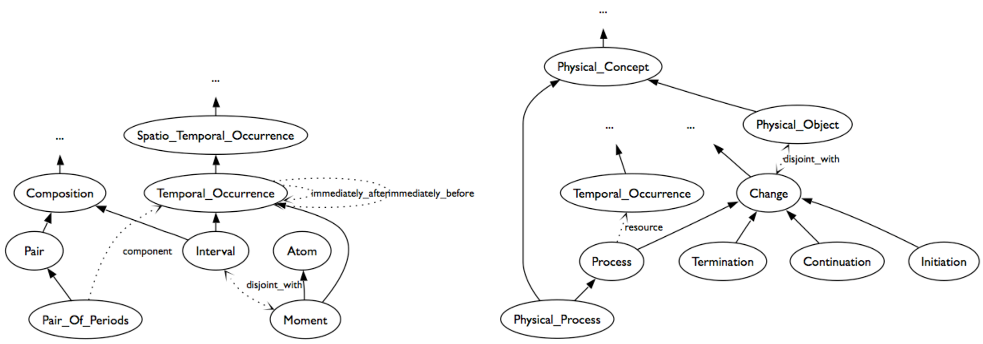

Figure 4. Concepts related to time and change.

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 `Process`es; they introduce causal propagation. Contrary to changes, processes are bound in time and space: they have duration and take place at a time and place. We furthermore distinguish `Physical_Process`es which operate on `Physical_Object`s. However, at this level we do not commit to a particular theory of causation or causal propagation.

4.2. The Intentional Level

Legal reasoning is based on a common sense model of intelligent behaviour, and the prediction and explanation of this intelligent behaviour. It is after all only behaviour of rational agents that can be effectively influenced by the law. The modules at the intentional level include concepts and relations necessary for describing this behaviour (i.e. `Action`s undertaken by `Agent`s in a particular `Role`) which are governed by law. Furthermore, it introduces concepts for describing the mental state of these agents, e.g. their `Intention` or `Belief`, but also communication between agents by means of `Expression`s.

The class of agents is defined as the set of things which can be the `actor` of an intentional action: they perform the action and are potentially liable for any effects caused by the action (see figure 5). `Action`s are processes, they are the changes performed by some agent who has the intention of bringing about the change. Because actions are processes they can become part of causal propagation, allowing us to reason backwards from effect to agent. Actions can be creative in that they initiate the coming into existence of some thing, or the converse. Also, actions are often a direct *re*action to some other action (see figure 5).

The agent is the medium of some intended outcome of the action: an action is always intentional. The intention held by the agent usually 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.

We distinguish between `person`s, individual agents such as “Joost Breuker” and “Pope Benedict XVI”, and `Organisation`s, aggregates of other organisations or per-

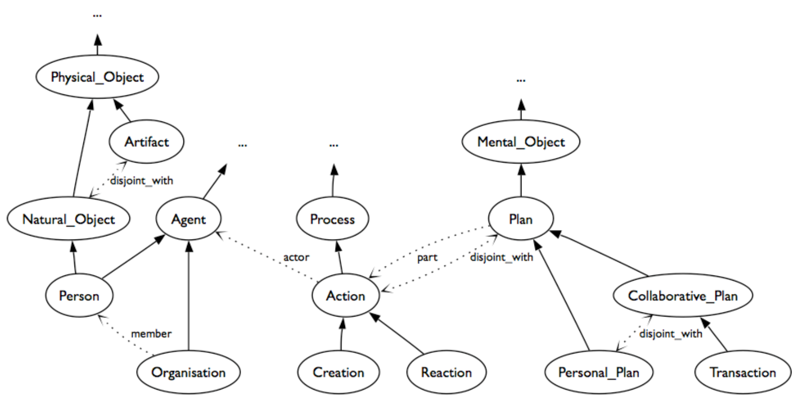

Figure 5. Actions, agents and organisations.

sons which acts ‘as one’, such as the “Dutch Government” and the “Sceptics Society”. `Artefact`s 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” and “Atlatl”[20]. Persons are physical objects as well, but are not designed (though some might hold the contrary) and are subsumed under the class of `Natural_Object`s. Note that natural objects can function as tools or weapons as well, the typical example being a stone, but are not designed for that specific purpose.

The notion of roles has played an important part in recent discussions on ontology [55,41,31]. 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. Roles specify standard or required properties and behaviour (see figure 6). 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 used as a model for intelligent decision making and planning[21]. 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.

[20] An atlatl is a tool that uses leverage to achieve greater velocity in spear-throwing, see `http://en.wikipedia.org/wiki/Atlatl`

[21] Regardless of whether it is a psychologically plausible account of decision making. Daniel Dennett’s notion of the *Intentional Stance* is interesting in this context (cf. [19]). Agents may do no more than occasionally apply the stance they adopt in assessing the actions of others to themselves.

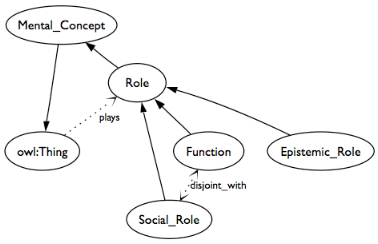

Figure 6. Roles.

The expression module covers a number of representational primitives necessary for dealing with Propositional_Attitudes (viz. [17]). 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 (cf. [53]). 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 aspect, the expression module is intentionally left under-defined. A rigourous definition of propositional attitudes relates them to a theory of reasoning and an argumentation theory. This argumentation theory is to be supplied by an argumentation ontology or framework. The theory of reasoning depends on the type of reasoning task (assessment, design, planning, diagnosis, etc.) LKIF is used in, and should be filled in (if necessary) by the user of LKIF .

Evaluative_Attitudes 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. I.e. 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 [22] and others. Of special interest is the Qualification, which is used to define norms based on [8]. 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).

4.3. 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

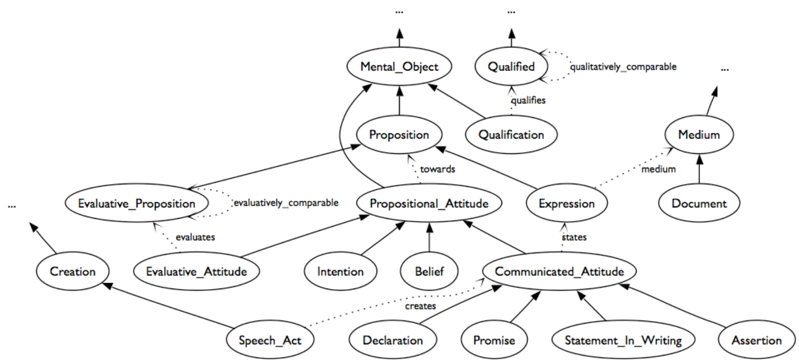

Figure 7. Propositions, Attitudes and Expressions.

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 [50,49]), typical legal roles, and concept definitions which allow us to express normative statements as defined in [8,6,7].

The Norm is a statement combining two performative 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 bringing about that the addressee brings about the more acceptable thing (cf. [48]), presumably through a sanction. These meanings do not have to occur together. 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).

A norm applies to (or qualifies) a certain situation (the Qualified situation), allows a certain situation – the Obliged situation or Allowed situation – and disallows a certain situation – the Prohibited or Disallowed situation, see Figure 8. The obliged and prohibited situation are both subsumed by the situation to which the norm applies. Besides that they by definition form a complete partition of the case to which the norm applies, i.e. all situation to which the norm applies are either a mandated case or a prohibited situation. This is true of the obligation and the prohibition: they are simply two different ways to put the same thing into words. The permission is different in that it allows something, but it does not prohibit anything. The logical complement of the mandated situation is here an opposite qualified situation, about which we know only that it cannot be obliged.

5. The Ontology in Practice: the Traffic Domain

The LKIF ontology not only provides a theoretical understanding of the legal domain, but its primary use in practice is as a tool to facilitate knowledge acquisition, exchange and representation: i.e. to formalise pieces of existing legislation. We evaluated the use of the ontology by formalising the EU Directive 2006/126 on driving licences,[22], a rela-

[22] The text is available on-line at http://eur-lex.europa.eu/.

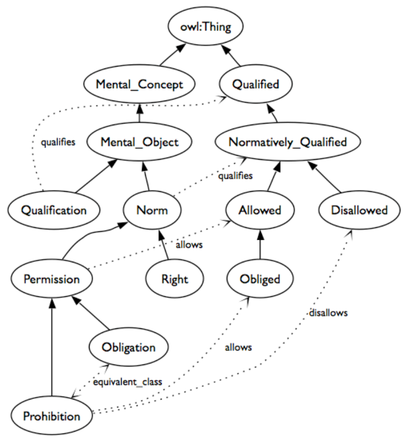

Figure 8. Qualifications and Norms

tively straightforward regulation, in which at least two types of normative statement are recognisable—definitional and deontic.

An example of a *definitional statement* from the EU directive is:

> Art. 4(2) **Category AM**: *Two-wheel vehicles or three-wheel vehicles with a maximum design speed of not more than 45 km/h.*

The mereo module of the ontology along with a qualified cardinality restriction (available with OWL 1.1) allows us to express that AM vehicles have two or three wheels:

$$\texttt{AM} \sqsubseteq 2\texttt{composed\_of.Wheel} \sqcup 3\texttt{composed\_of.Wheel}.$$

Modelling 'design speed not more than 45 km/h' is more challenging as it requires us to represent the rather common sense domain of speeds, distances etc. Of course, one could introduce the datatype property `designSpeed` and require its value be expressed in km/h.This choice, however, would not make justice of the conceptual complexity involved in 'design speed not more than 45 km/h', which contains reference to several notions: unit of measurement, number, designed speed, and a no-more relation. In fact, 'design speed not more than 45 km/h' can be rendered by imposing an *linear ordering* relation `less-than` on the different (instances of the) subclasses of the class `DesignSpeed`.[23] The ordering allows us to define the class of those `DesignSpeed`s with a value not exceeding some `N45`—i.e., $\forall$`less-than.DesSpeed-km-h-45`.

Let us now look at an example of a *deontic statement*:

> Art. 4(2) *The minimum age for category AM is fixed at 16 years.*

Art. 4(2) expresses an obligation whose logical form can be rendered by the implication:

If x is driving a AM vehicle, then x **must** *be at least 16 years old*.

---

[23] The ordering is linear—i.e, reflexive, antisymmetric, transitive and total—since it mirrors the ordering of the natural numbers. For whenever $n \leq m$, we have that `DesignSpeed-km-h-n`(a) `less-than` `DesignSpeed-km-h-m`(b), with a, b instances.

To fix some terminology, the antecedent is the *context* to which the obligation applies; the consequent (minus the deontic operator **must**) is the *content* of the obligation itself (what the obligations prescribes it ought to be the case). Consistently with this analysis, the LKIF ontology defines obligations as classes (see Section 4.3).

In our case, art. 4(2) allows the situation `DriverAM` $\sqcap$ `DriverOlderThan16` and forbids `DriverAM` $\sqcap \neg$`DriverOlderThan16`. Suppose that the classes `DriverOlderThan16` and `DriverAM` have already been defined.[24] To model the obligation that drives of AM vehicles must be at least the 16 years older, we introduced the obligation-type class `MinAgeAM` as follows:

$$\begin{aligned}
\texttt{MinAgeAM} &\sqsubseteq \forall\texttt{allows}.(\texttt{DriverAM} \sqcap \texttt{DriverOlderThan16}).\\
\texttt{MinAgeAM} &\sqsubseteq \exists\texttt{allows}.(\texttt{DriverAM} \sqcap \texttt{DriverOlderThan16}).\\
\texttt{MinAgeAM} &\sqsubseteq \forall\texttt{disallows}.(\texttt{DriverAM} \sqcap \neg\texttt{DriverOlderThan16}).\\
\texttt{MinAgeAM} &\sqsubseteq \exists\texttt{disallows}.(\texttt{DriverAM} \sqcap \neg\texttt{DriverOlderThan16}).
\end{aligned}$$

Other deontic operators, such as permission or prohibition, can be accounted in an alike manner (see [7]). Notwithstanding the parsimony of this type of definition, using the LKIF ontology to model normative statements proves to be rather straightforward. Of course, a specialised modelling environment for legislative drafters would need to provide a shorthand for such standard OWL definitions.[25]

The representation of art. 4(2) suggests the LKIF ontology be augmented with a module taking care of quantities, units of measurement, numbers, fractions, mathematical operations, and the like. This is crucial not only for the EU Directive 2006/126, in which most definitional statements contain quantitative features of vehicles (e.g., power, cylinder capacity); quantities and calculations play a central role in any legislative text. Note, however, that the LKIF ontology can only provide a *purely terminological* account, without being able to do mathematical computations. This is unavoidable, given that OWL is a purely logical language. We are currently investigating whether we can import an existing OWL ontology dealing with measurements, such as PhysSys/SUMO or from the Ontolingua server[26].

6. Discussion

We 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 initial requirements for this 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. We discussed several other ontologies, argued where they fall short in providing this grounding, and presented an approach based on insights from cognitive science

[24] The class `DriverOlderThan16` can be defined by using a `more-than` ordering relation, roughly along the same lines as the class $\forall$`less-than.DesSpeed-km-h-45`. The class `DriverAM` can be easily defined.

[25] See e.g. the SEAL project, `http://www.leibnizcenter.org/project/current-projects/seal`

[26] See `http://www.ksl.stanford.edu/software/ontolingua/`

and knowledge engineering, rather than philosophy. This approach was further discussed as the basis for a comprehensive methodology for ontology development, including a principled distinction between frameworks and ontologies. The LKIF Core ontology was developed using this methodology, and is composed of fourteen modules, two of which represent frameworks. We briefly introduced the main concepts in these modules and gave an example of how the ontology can be used in modelling an EU directive.

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 in the previous phases (See [13]). 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 ensures the ability to infer the type of individuals; this does not hold for partial class definitions (using only necessary conditions).

In retrospect, this convention turned out to pose severe problems for existing OWL DL and OWL 1.1 reasoners as their performance is significantly affected by the generic concept inclusion axioms (GCI): axioms where the left-hand side of a `rdfs:subClassOf` statement is a complex class definition. These axioms are abundant when defining classes as equivalent to e.g. `owl:someValuesFrom` restrictions and in combination with lots of inverse property definitions, this creates a large completion graph for DL reasoners[27] As result of these findings, the LKIF ontology has underwent a significant revision since its initial release. However, reasoner performance remains an issue with large, complex ontologies such as LKIF Core.

Using LKIF Core in practice, as e.g. in the traffic example, points to the traditional knowledge-acquisition bottle-neck: for any formal representation of any domain, one needs to build formal representations of adjoining domains. As has been said, this can be largely overcome by including specialised foundational or domain ontologies in a representation based on the LKIF ontology provided that the quality of these ontologies is sufficient. Depending on availability we might consider providing a library of 'compatible' ontologies to users of LKIF Core. This will be of especial use when the ontology vocabulary is adopted for expressing the LKIF vendor models that are being developed within the Estrella project. Recent experiences suggest that the ontology should be extended with not only a module for expressing measurements, but also frameworks that allow us to adequately capture transactions and other more complex action structures.

With respect to coverage of the legal domain, the purpose of the study outlined in Section 2.1 is more ambitious than only the selection of the most basic terms for describing law, but time and effort constraints make it that we could only consider a small pool of referents. The list of terms will be subjected to a more rigourous empirical study, whereby we will consult a group of legal professionals (taking courses in legal drafting), and law students. These empirical studies are planned in the sideline of Estrella. By applying statistical cluster analysis, we might be able to identify the properties of the scales used (e.g. are they independent?) and whether the statistical clusters have some resemblance to the clusters we have identified based on theoretical considerations. The results of this analysis will be used to evaluate the ontology compared to the requirements we identified in the previous chapters.

[27] Thanks to Taowei David Wang for pointing this out, see `http://lists.owldl.com/pipermail/pellet-users/2007-February/001263.html`

The LKIF ontology is available online as separate but interdependent OWL DL files, and can be obtained from the Estrella website at `http://www.estrellaproject.org/lkif-core`. This website also provides links to online documentation and relevant literature.

References

[1] J. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23:123–154, 1984.

[2] James F. Allen and George Ferguson. Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4(5):531–579, 1994.

[3] S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D. L. McGuinness, P. F. Patel-Schneider, and L. A. Stein. OWL web ontology language reference. W3C recommendation, World Wide Web Consortium, February 2004. M. Dean, G. Schreiber (eds.).

[4] Palash Bera and Yair Wand. Analyzing OWL using a philosophy-based ontology. In Achille C. Varzi and Laure Vieu, editors, *Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004)*, volume 114 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2004.

[5] Olivier Bodenreider, Barry Smith, and Anita Burgun. The ontology-epistemology divide: a case study in medical terminology. In Acchille Varzi and Laure Vieu, editors, *Formal ontology in Information Systems*, pages 185–198. IOS-Press, Amsterdam, 2004.

[6] Alexander Boer. Note on production rules and the legal knowledge interchange format. Technical report, Leibniz Center for Law, Faculty of Law, University of Amsterdam, 2006.

[7] Alexander Boer, Thomas F. Gordon, Kasper van den Berg, Marcello Di Bello, András Förhécz, and Réka Vas. Specification of the legal knowledge interchange format. Deliverable 1.1, Estrella, 2007.

[8] Alexander Boer, Tom van Engers, and Radboud Winkels. Mixing legal and non-legal norms. In M-F. Moens and P. Spyns, editors, *Jurix 2005: The Eighteenth Annual Conference.*, Legal Knowledge and Information Systems, pages 25–36, Amsterdam, 2005. IOS Press.

[9] J.A. Breuker and W. Van de Velde, editors. Common*KADS Library for Expertise Modelling*. IOS Pres, 1994.

[10] Joost Breuker, Alexander Boer, Rinke Hoekstra, and Kasper van den Berg. Developing content for LKIF: Ontologies and frameworks for legal reasoning. In Tom M. van Engers, editor, *Legal Knowledge and Information Systems. JURIX 2006: The Nineteenth Annual Conference*, volume 152 of *Frontiers in Artificial Intelligence and Applications*, 2006.

[11] Joost Breuker and Rinke Hoekstra. Core concepts of law: taking common-sense seriously. In A.C. Varzi and L. Vieu, editors, *Proceedings of Formal Ontologies in Information Systems (FOIS-2004*, pages 210–221. IOS-Press, 2004.

[12] Joost Breuker and Rinke Hoekstra. Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law. In *Proceedings of EKAW Workshop on Core ontologies*. CEUR, 2004.

[13] Joost Breuker, Rinke Hoekstra, Alexander Boer, Kasper van den Berg, Rossella Rubino, Giovanni Sartor, Monica Palmirani, Adam Wyner, and Trevor Bench-Capon. OWL ontology of basic legal concepts (LKIF-Core). Deliverable 1.4, Estrella, 2007.

[14] Joost Breuker, Andre Valente, and Radboud Winkels. Use and reuse of legal ontologies in knowledge engineering and information management. In V.R. Benjamins, P. Casanovas, J. Breuker, and A.Gangemi, editors, *Law and the Semantic Web*, pages 36 – 64. Springer, 2004.

[15] Pompeu Casanovas, Nuria Casellas, Joan-Josep Vallbe, maria Poblet, Richard Benjamins, Mercedes Blazquez, Raul Pena, and Jesus Contreras. Semantic web: a legal case study. In Jahn Davies, Rudi Studer, and Paul Warren, editors, *Semantic Web Technologies*. Wiley, 2006.

[16] B. Chandrasekaran and Todd R. Johnson. Generic tasks and task structures: History, critique and new directions,. In J. M. David, J. P. Krivine, and R. Simmons, editors, *Second Generation Expert Systems*. Springer Verlag, 1993.

[17] Mats Dahllöf. On the semantics of propositional attitude reports, 1995.

[18] R. Davis, H. Shrobe, and P. Szolovits. What is a knowledge representation? *AI Magazine*, Spring:17–33, 1993.

[19] D. Dennett. *The Intentional Stance*. MIT-Press, 1987.

[20] M. Donnelly. Relative places. *Applied Ontology*, 1(1):55–75, 2005.

[21] Mariano Fernández, Asunción Gómez-Pérez, and Natalia Juristo. Methontology: from ontological art towards ontological engineering. In *Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering*, pages 33–40, Stanford, USA, March 1997.

[22] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening ontologies with DOLCE. In A. Gomez-Perez and V.R. Benjamins, editors, *Proceedings of the EKAW-2002*, pages 166–181. Springer, 2002.

[23] A. Gangemi and P. Mika. Understanding the semantic web through descriptions and situations. In *Proceedings of CoopIS/DOA/ODBASE*, pages 689–706, 2003.

[24] A. Gangemi, M.T. Sagri, and D. Tiscornia. A constructive framework for legal ontologies. In V.R. Benjamins, P. Casanovas, J. Breuker, and A. Gangemi, editors, *Law and the Semantic Web*, pages 97–124. Springer Verlag, 2005.

[25] M.R. Genesereth and R.E. Fikes. Knowledge interchange format, version 3.0 reference manual. Technical Report Logic-92-1, Computer Science Department, Stanford University, June 1992.

[26] T.F. Gordon, H. Prakken, and D. Walton. The carneades model of argument and burden of proof. *Artificial Intelligence*, 171(10-11):875–896, 2007.

[27] T. R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.

[28] T. R. Gruber. Toward principles for the design of ontologies used for knowledge sharing. In N. Guarino and R. Poli, editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Deventer, The Netherlands, 1994. Kluwer Academic Publishers.

[29] Michael Grüninger and Mark S. Fox. Methodology for the design and evaluation of ontologies. In *IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing*, 1995.

[30] N. Guarino and P. Giaretta. Ontologies and knowledge bases: Towards a terminological clarification. In N.J. Mars, editor, *Towards Very Large Knowledge Bases – Knowledge Building and Knowledge Sharing 1995*, pages 25–32, Amsterdam, 1995. IOS Press.

[31] N. Guarino and C. Welty. Evaluating ontological decisions with OntoClean. *Communications of the ACM*, 45(2):61–65, 2002.

[32] Nicola Guarino. The ontological level. In R. Casati, B. Smith, and G. White, editors, *Philosophy and the Cognitive Sciences*. Hölder-Pichler- Tempsky, Vienna, 1994.

[33] P. J. Hayes. The second naive physics manifesto. In J. R. Hobbs and R. C. Moore, editors, *Formal Theories of the Common Sense World*, pages 1–36. Ablex Publishing Corporation, Norwood, 1985.

[34] Rinke Hoekstra, Joost Breuker, Marcello Di Bello, and Alexander Boer. The LKIF Core ontology of basic legal concepts. In Pompeu Casanovas, Maria Angela Biasiotti, Enrico Francesconi, and Maria Teresa Sagri, editors, *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, June 2007.

[35] George Lakoff. *Women, Fire and Dangerous Things*. University of Chicago Press, 1987.

[36] George Lakoff and Rafael Núñez. *Where Mathematics Comes From*. Basic Books, 2000.

[37] Guiraude Lame. Using NLP techniques to identify legal ontology components: Concepts and relations. *Artificial Intelligence and Law*, 2006. this issue.

[38] Douglas B. Lenat. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.

[39] Sandra Marcus, editor. *Automating Knowledge Acquisition for Expert Systems*. Kluwer, Reading MA, 1988.

[40] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. WonderWeb ontology library. Deliverable D18, version 1, ISTC-CNR (Italy), 2003.

[41] C. Masolo, L. Vieu, E. Bottazzi, C. Catenacci, R. Ferrario, Aldo Gangemi, and Nicola Guarino. Social roles and their descriptions. In *Proceedings of Knowledge Representation Workshop*, 2004.

[42] John McCarthy. Programs with common sense. In *Proceedings of the Teddington Conference on the Mechanization of Thought Processes*, pages 75–91, London, 1959. Her Majesty's Stationary Office.

[43] M. Minsky. A framework for representing knowledge. In P. H. Winston, editor, *The psychology of Computer Vision*, pages 211–277, New York, 1975. McGraw-Hill.

[44] Marvin Minsky. Why people think computers can't. *AI Magazine*, 3(4), Fall 1982.

[45] Marvin Minsky. *The Society of Mind*. Simon & Schuster Inc., 1984.

[46] Enrico Motta. *Reusable Components for Knowledge Modelling*. FAIA-Series. IOS Pres, Amsterdam

NL, 1999.

[47] I. Niles and A. Pease. Towards a standard upper ontology. In Chris Welty and Barry Smith, editors, *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)*, Ogunquit, Maine, October 17-19 2001.

[48] Jan Nuyts, Pieter Byloo, and Janneke Diepeveen. On deontic modality, directivity, and mood, 2005.

[49] Rosella Rubino, Antonino Rotolo, and Giovanni Sartor. An owl ontology of fundamental legal concepts. In Tom M. van Engers, editor, *Legal Knowledge and Information Systems. JURIX 2006: The Nineteenth Annual Conference*, volume 152 of *Frontiers of Artificial Intelligence and Applications*. IOS Press, 2006.

[50] Giovanni Sartor. Fundamental legal concepts: A formal and teleological characterisation. Technical report, European University Institute, Florence / Cirsfid, University of Bologna, 2006.

[51] R. Schank and R. Abelson. *Scripts, Plans Goals and Understanding*. Lawrence Erlbaum, New Jersey, 1977.

[52] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. Van den Velde, and B. Wielinga. *Knowledge Engineering and Managament: The CommonKADS Methodology*. MIT Press, 2000.

[53] John Searle and Daniel Vanderveken. *Foundations of Illocutionary Logic*. Cambridge University Press, Cambridge, 1985.

[54] John F. Sowa. *Knowledge Representation: Logical Philosophical, and Computational Foundations*. Brooks Cole Publishing Co, Pacific Grove, CA, 2000.

[55] F. Steimann. On the representation of roles in object-oriented and conceptual modelling. *Data and Knowledge Engineering*, 35:83–106, 2000.

[56] Mike Uschold and Michael Grüninger. Ontologies: principles, methods, and applications. *Knowledge Engineering Review*, 11(2):93–155, 1996.

[57] Mike Uschold and Martin King. Towards a methodology for building ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95*, Montreal, Canada, 1995.

[58] A. Valente. *Legal knowledge engineering: A modelling approach*. PhD thesis, University of Amsterdam, 1995.