Use and Reuse of Legal Ontologies in Knowledge Engineering and Information Management

Joost Breuker[1], André Valente[2], and Radboud Winkels[1]

[1] University of Amsterdam, Leibniz Center for Law,
Faculy of Law, P.O. Box 1030, NL 1000 BA Amsterdam, Netherlands
{breuker, winkels}@lri.jur.uva.nl
[2] Knowledge Systems Ventures,
7707 Boeing Ave, Los Angeles, CA 90045, USA
avalente@ksventures.com

Abstract. In this article we present an overview of legal ontological modeling over a period of more than a decade. Most of the research reported concerns results from mid-size (European) projects aimed at the development of legal reasoning and information management tools and systems. In these projects we developed ontologies for several legal or regulation domains. However, the main thread of this article is provided by fundamental research performed by us or under our supervision (e.g. PhD theses by [1], [2], [3]), leading to more abstract legal 'core' ontologies and legal reasoning architectures. The major insights we have obtained from these experiences can be summarized as follows:

- Legal sources contain and assume non-legal, common-sense based domain knowledge. Therefore, a legal core ontology should be rooted in highly abstract common sense concepts as a foundational ontology. This notion is worked out in the *LRI-Core ontology*, presented in this article.
- What remains as typical legal knowledge to be modelled is normative and legal responsability knowledge. However, in this article we only summarize our work in this area, refering to [4] and to [5] on these issues.
- As law and legal theory is focussed on questions of justification, legal (core) ontologies, e.g. *FOLaw*, are epistemological frameworks, describing legal reasoning, rather than legal ontologies, explaining knowledge resources.

1 Introduction

This article presents insights acquired during more than a decade of development, use and reuse of legal ontologies at the Leibniz Center for Law (LRI) at the University of Amsterdam. These ontologies have been constructed in various projects concerned with the development of legal knowledge systems and legal information management. These insights provide a framework rather than a methodology for modeling legal knowledge and reasoning. Results and experiences in the projects are used to illustrate this framework. Although this framework is operationalized by reusable legal knowledge system architectures (e.g. *ON-LINE* [4], TRACS [6], [2]), the focus of this article is on discussing the conceptual views on legal knowledge and reasoning.

V.R. Benjamins et al. (Eds.): Law and the Semantic Web, LNAI 3369, pp. 36–64, 2005.
© Springer-Verlag Berlin Heidelberg 2005

In the projects a large variety of legal domains have been analyzed and modeled: traffic regulations; tax –, criminal –, and administrative law; international treaties on trade, and safety at sea. To enable reuse of these ontologies in a library and to abstract the common legal denominators of these legal domains, much effort has been spent in finding a unifying view on legal domains. The goal is to support the modeling of ontologies in new legal domains. In this paper we will discuss two proposals for a legal **core ontology**: FOLaw [4] and more recently: LRI-Core. FOLaw has proven to be a good modeling support for building legal knowledge systems, because it reflects an understanding of types of knowledge and dependencies in legal reasoning. However, when applied to ontology based management of legal information services, the support is very limited: in Section 5 we will explain why and what views have lead us to develop this new LRI-Core ontology.

An ontology describes how some domain is 'committed' to a particular view: not so much by the collection of the terms involved but in particular by the way these terms are structured and defined. This structure tells us "what a domain is about". For instance, medical domains are about malfunctions of humans. These malfunctions are often diseases, i.e. processes; they are classified in (multiple) taxonomies, and associated with sets of typical symptoms, and with treatments ([7, 8]; more recently: [9]). In the next section (Section 2) we will make such an analysis for legal domains.

The article is further structured as follows. We will give a short overview of ontological commitments that can be found in legal theory. Next we will introduce our 'functional ontology of law' and see that it is rather an epistemic framework than a detailed ontology. From there on we will explore what is in a true ontology of law; what are typical and unique elements of law? When we have identified them, we will see them put to use in several applied R&D projects. Finally, we will discuss the current state of our work and draw some conclusions.

2 What Is Law About?

An ontology makes explicit the views one is committed to in modeling a domain. Modeling is taken here in the broad sense that includes the notion of understanding. A major and typical problem from jurisprudence (legal theory) occurs already in the use of the term "law". In the title of this section we avoided any commitment to whether we mean '*the* law' or 'laws'. Indeed, the problem of what counts as the unit of law is already one of the fundamental questions in jurisprudence and is called the *individuation* problem:

> "Classifying laws in logically distinct categories has always been one of the major tasks of legal philosophy ... The classification of laws presupposes a solution to the more fundamental problem of the individuation of laws, i.e., an answer to the question 'What is to count as one complete law?" [10–page 825]

There are two extreme views. The first one takes all legally valid statements in legal sources (legislation, precedence law, etc) as a whole: the law. The assumption is that in principle the individual statements constitute a coherent whole. This is the predominant view in jurisprudence and legal philosophy (see also Section 3). Whether this coherence is an actual concern for the legal system (i.e., the law should be the object of proper knowledge management), or whether it is 'genetically' built-in by the constraints pro-

vided by fundamental, 'natural' legal principles, is a long and classical debate in legal theory.

However, the other extreme takes all legally valid statements as being individual laws. *In extremo* this view is incorrect, if only because at least one other legal (meta-)statement that is concerned with the validity of other legal statements: we need law about law, and this dependency is a serious concern of legal scholars (see e.g. [11]).

Legal theory is in the first place concerned with justifying law, so legal scholars will not easily take validity statements in law as a side issue

Law is theoretically very much concerned with the 'version' control of legislation, but in practice the means to control which is the current, valid version are not very sophisticated and not foolproof; let alone to assess whether some legal statement was valid at some point in time in the past. In constructing legal reasoning or information systems, knowledge engineers simply assume that the legal source to be modelled is a legally valid one. In the rest of this article we leave this meta-level issue in legal reasoning aside (see [12] further on this issue).

The coherence of law in a legal system as postulated by many legal scholars is in the first one that should evade contradictory outcomes for assessing legal cases: the law should not contain serious contradictions (exceptions are not considered to be contradictions; see Section 2.1). However, we usually mean by coherence more than only 'not contradictory': there should also be no conceptual 'gaps': two unrelated statements do not contradict one another but they do not form a coherent text. This lack of 'semantic coherence' [13] is exactly the problem one will find in modelling legislation. The normative statements that make up the bulk of legislation refer to some domain of social activity, but do not describe this domain. The individual normative statements qualify certain kinds of situations as illegal. By another meta-level law that states that anything that is not forbidden by law is allowed – the default of all national legal systems – in principle no statements are required that fill the set of situations that are not illegal. A legal source is therefore incomplete and It means that a knowledge engineer has to reconstruct and reverse engineer the social domain some legal source is about. In the next section we will illustrate this principle for the domain of traffic.

2.1 Distinguishing Normative and World Knowledge: TRACS

"Can you develop a computer program that can check if the new traffic regulation (RVV-90) is consistent and complete?" This apparently innocent question, posed by a government agency concerned with traffic safety, SWOV [1], triggered a decade of research at our institute. The question is highly similar to the verification of software.

Different from a 'normal' text or a computer program, the individual articles have no semantic coherence. The individual articles in the RVV-90 refer to situations in traffic, but there is no (discourse) structure that connects the statements. Therefore, it will be hard to assess whether a regulation is complete. Viewed as a text, the RVV-90, and for that matter (almost) any regulation, is full of gaps as it is not intended that one statement can and should be related to a next one. This kind of incompleteness can only

[1] Stichting Wetenschappelijk Onderzoek Verkeersveiligheid; Foundation for Research on Traffic Safety.

be detected when one has reconstructed the model of the domain the legislator had in mind when drafting a regulation. Even then it is hard to assess whether the legislation covers all situations intended, because the intentions of a law are only in an abstract and incomplete way specified.

One may, for instance, observe that the RVV-90 does not talk about children, while they are the most endangered species in the traffic jungle. The question is whether these kinds of gaps are accidental (and non-intended), or due to the limited role of law, or whether there simply is no real gap. One may argue that children have insufficient control or understanding to 'obey' the law. Also: they are subsumed (implied) in the RVV-90 under pedestrian, etc. There are indirect ways to establish this covering or completeness. The first one is to use constraints (requirements, goals). One of the aims of the RVV-90 is to avoid collisions, i.e. if there is a collision, at least one participant should have trespassed at least one norm. The other one is to compare the same situation under two different codifications. For instance, the legislator has indicated in which respects the RVV-90 is aimed to be different (an improvement) from the previous version. At the end of this section we will elaborate on this issue.

If it is difficult to assess the completeness of a regulation, consistency is another problem, as can be illustrated by the first normative statements in the RVV-90:

Art 3.1 Vehicles should keep as much as possible to the right
Art 3.2 Two bicyclists may ride next to each other

Article 3.1 describes an obligation and 3.2 a permission to a subset of drivers: the drivers of bicycles. In fact, it is not so easy to see how 3.2 is an exception to 3.1. It requires some complex spatial reasoning ('right', 'next to') to see that the left-hand bicyclist violates article 3.1, because she leaves the right bicyclist between her and the side of the lane. [2] This exception is a normative conflict and therefore a logical inconsistency. This inconsistency can be repaired by applying meta-rules, such as the principles that provide priority to the more recent, the higher and the more specific norm.

Exceptions are intended. They are used to limit over-generalizations of more generic norms, in the same way as we still want to classify penguins under birds despite the fact that they lack one of the essential characteristics of birds: that they can fly. [3] There is a way

[2] In fact, the exception is even more subtle because the left-hand bicyclist is also held to keep still as much as possible to the right, i.e. the exception is not a license for this bicyclist to take any position in the lane next to the right-hand bicyclist.

[3] This is the typical example to illustrate the need for non-monotonic reasoning in normative (deontic) reasoning. However, this analogy between exceptional birds and normative exceptions does not necessarily imply that normative reasoning with exceptions is really non-monotonic. The conflict resolution, implied by the principle that a more specific normative statement overrides a general one, suggests so, but a closer look at what the conflict resolution means is that there is no retraction of beliefs of what is the case in the world. The case remains unchanged. It is only that there are conflicting normative qualifications about the case. The typical view in jurisprudence and AI & Law is that the conflict is resolved by putting priorities to applicable, conflicting norms: not to withdraw norms or facts about the case. Article 3.a is still valid for the left-hand bicyclist: she still has to keep as much as possible to the right. Despite this observation, most researchers assume that normative reasoning is non-monotonic because one has to 'withdraw' some conclusion (one of the two qualifications).

to get rid of normative exceptions without affecting the normative qualifications intended by the code [14]. We may be able to see exceptions as "not really inconsistencies" in law, but the problem remains that we cannot distinguish on formal grounds between intended and non-intended exceptions. In fact in the TRACS project [15, 16] we have found out that the RVV-90 contains many non-intended exceptions which are often hidden in the (implicit) normative structure of a regulation. We should add that normative conflicts, or more precise: "conflicts of disaffirmation" are not the only kind of inconsistencies that may occur in regulations. [17] also distinguishes "compliance conflicts". These conflicts are between mandatory norms (obligations and prohibitions) which are not jointly realizable. For instance, a bus driver may be obliged to keep the time-table, but at the same time speed-limits may prohibit him to be able to comply with the time-table norm.

TRACS is a prototype knowledge system aimed at verifying the Dutch traffic regulation RVV-90. TRACS is a kind of policeman that assesses all possible traffic situations the RVV-90 may distinguish. The situation generator (or recognizer) constructs a a combination of traffic participants and traffic actions on some configuration of roads. These situations are generated by an ontology that represents the definitions and axioms of the terms involved. This 'world' knowledge base allows one to infer, for instance, all possible spatial relations: e.g.

```
equivalent(left-of(car-A,bicycle-B),right-of(bicycle-B,car-A))
```

These were used to match with the set of regulations, coded in terms of the traffic ontology so that the legal consequences could be derived. The RVV-90 was never completely tested but in a long series of test trials many errors were found: the most notable one was the fact that the tram was not allowed to run on the tram-way (see [2] for all details). The essential issue here is that this ontology of traffic terms plus an ontology about spatial terms was sufficient to test the core of the traffic regulation. The major categories of this ontology were roles (drivers), vehicles, actions, and parts-of-road. The spatial ontology is a very simple, abstract one and consisted of axioms of terms of orthogonal directions and positions in a two dimensional world.

3 Ontological Commitments in Legal Theory

Legal theories usually contain elements of an ontology, but they are normally framed under some specific theoretical goal which lies beyond the ontology itself. For instance, Hart's theory intends to explain how legal systems evolve; Kelsen's goal was to demonstrate the difference between laws and morals; yet both propose very specific views on what competence is behind legal phenomena, and what primary concepts are used to represent law.

Kelsen. In his last work ('General Theory of Norms' [11]), Kelsen proposed four basic types of norms: command, empower, permit and derogate. Commanding norms command (prohibit, obligate) a certain behavior. Empowering norms associate roles with the power to posit and apply norms under certain restrictions. Permitting norms refer to what he called the positive sense of permission. Kelsen argued that we may permit behavior in the sense that this behavior is neither prohibited nor commanded, in which case we have a negative or *weak*) permission. In contrast, permitting norms use a positive sense

of permission, in which behavior is actively allowed. Kelsen sees permitting norms as always an exception to a command, i.e. it occurs when there is a commanding norm about a certain behavior, and this command is then derogated by the permitting norm. Kelsen points out that the difference between commanding and prohibiting on the one hand, and permitting (positively) and empowering on the other, can be mapped onto the distinction between observing/violating and applying a norm. Only a norm which commands (or prohibits) a certain behavior can be observed or violated. Norms that permit or empower cannot be violated or observed; only applied or not. Finally, derogating norms "repeal of the validity of [. . .] another norm". Kelsen stresses that a derogating norm does not repeal another norm, but its validity: a derogated norm still exists, but it is no longer valid.

Hart. The Hartian distinction between *primary* and *secondary rules* (norms) has become a quasi-standard in legal theory. Hart's distinction, carefully detailed in his 'Concept of Law' [18], draws a line between a first level which refers to human behavior and a second, meta-level of the first, which contains knowledge *about* primary norms. These secondary rules may belong to three types: (i) rules of adjudication, that can be used to determine authoritatively whether a certain primary rule has been violated or not; (ii) rules of recognition which define, directly or indirectly, which rules are the valid ones, and can therefore be applied; (iii) rules of change, which define how rules are to be made, removed or changed. These distinctions point out three functions of secondary norms: to provide support for solving conflicts (adjudication), to specify the limits of the legal system (recognition) and to specify how the legal system can change in time (change).

Bentham. Bentham's theory is divided in two parts [19]. The first is a logic of imperation which uses four basic operators: commanded, prohibited, non-commanded and permitted.[4] These are in fact interdefined, resulting in only one of the four as primitive. The second part is a logic of obligations and rights in which he defines three primitive concepts: obligation, right to a service and liberty. These are also interdefined based on obligation, which Bentham sees as an obligation someone has to the effect that something (some state of affairs) occurs. Therefore, we are left with only two atomic concepts: commanded and obligation.

Hohfeld. Hohfeld's theory is considered a landmark in American jurisprudence [22]. An interesting (and unusual) aspect of Hohfeld's theory is that rights and other positional concepts that represent legal relations are considered primitives. There are two groups of interrelated legal relations or positions. The first group is composed by right, duty, no-right, privilege and has a strong normative flavor. These concepts are closely related to Bentham's concept of right, obligation and liberty. The second group consists of power, liability, disability and immunity. These concepts are more closely related to legal competences and legal responsibilities.

3.1 Legal Theory in AI and Law

There are a number of studies in AI & Law which have used ontological assumptions drawn from legal theory in the manner we propose in this article. Allen and Saxon

[4] A logic of imperation — an idea also mentioned by Austin [20] — was later developed in more detail by Hofstadter [21], but it is presently considered to be superseded by deontic logics.

(e.g. [23, 24]) have developed a 'language for legal relations' (LLR) in which they have transformed the Hohfeldian ontological primitives into about forty relations ('cascading propositions'), taking the notion of duty as primitive. Hamfelt [25] proposed and implemented a representation of legal knowledge in which Hart's primary and secondary rules were mirrored in meta-levels of a logic programming formalism. [1] has formalized normative legal knowledge mainly on the basis of legal theoretical concepts from Kelsen and Hart. This is also the perspective on legal normative knowledge we take here.

There is also work which has an ontological flavor, but which has not been based on legal theory. For example, McCarty's Language of Legal Discourse [26] can be seen as an ontology of Law, where his 'modalities' play the role of knowledge categories and are linked together with a formal (logical) presentation. Also, the research in deontic logics as a basis for normative reasoning sometimes uses ontological assumptions from or is applied to legal theory — see for instance [27, 28]. Kelsen's view on norms as descriptions of an ideal world can be seen the basis of deontic logic. Deontic logic provides interpretations for the terms 'obligation', 'prohibition' and 'permission'. Indeed some form of deontic logic is often proposed as a formalism for automated normative reasoning. Because (standard) deontic logic is intractable, and gives rise to pseudo-paradoxes, all kind of extensions and simplifications have been developed. In fact, [1] has shown that deontic logic does not make the necessary distinction between the normative status of *a* situation and the normative operator of a norm. By making this distinction a much simpler and tractable inference mechanism has been formally defined and implemented. Van Kralingen ([29]) uses the theory of Brouwer ([30]) as a starting point for what he sees as the core elements of law: norms, actions and (legal) concepts.

From this overview of legal theoretical studies we conclude that for over a century the main interest has been on the normative aspects of legal knowledge. [5] These studies deal with concepts like rights, permissions, obligations, etc. and their interrelations. However, this ontological perspective on the normative core business of law has resulted rather in epistemological views on legal reasoning, as exemplified in studies in deontic logics (e.g. [32]), than in a comprehensive view of what (categories of) concepts make up law (see also [33] for the way ontological and epistemological views are intertwined in views on normative knowledge.)

4 Knowledge Typing and Dependencies in Legal Reasoning: *ON-LINE* and *FOLaw*

Although the combination of world and normative knowledge makes up the resource for reasoning in legal domains, this is only part of the story. In the mid-90ies, André Valente constructed a core ontology that distinguished also other types of knowledge [1, 4]. This core-ontology, called *FOLaw*, served a number of purposes. The first one was to distinguish the various types of knowledge in legal reasoning, and in particular those types that are typical for legal reasoning. Related to this role it also explained the dependencies between these types of knowledge in legal reasoning. This typing and its dependencies could easily be translated into an architecture for legal reasoning: *ON-LINE*. The sec-

[5] A notable exception is the work by [31] on legal causation (see also [5] in this volume.

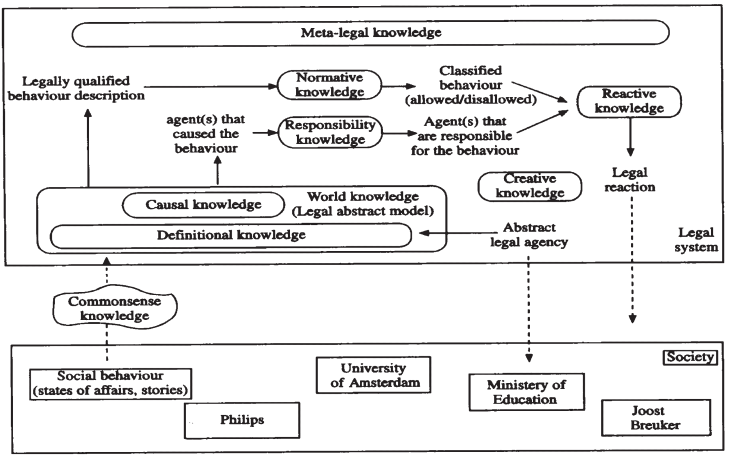

Fig. 1. FOLaw: a functional ontology for law [Valente, 1995]

ond role is the typical core-ontology role: to organize and index libraries of domain ontologies and to support knowledge acquisition to construct new ontologies.

The ontology is a functional ontology. This means that the roles the legal system plays in society are taken as point of departure. A legal-sociological view is taken, rather than a perspective from the law itself, as in most legal theoretical studies. There is a secondary notion of functional involved: *FOLaw* identifies the dependencies between the types of knowledge, which indicate the roles that types of knowledge play in the reasoning. These two views on 'functional' are not independent. One may see the reasoning as to some extent simulating the social roles, in the same way as reasoning about physical systems consists to a large extent of simulating physical processes.

We will give here a summary description of *FOLaw*. Figure 1 provides the comprehensive picture of dependencies of the various types of knowledge in legal reasoning. At the same time it also expresses the role of the legal system as controlling the actual social behavior of individuals and organizations in society.

4.1 *FOLaw*'s Types of Knowledge in Legal Reasoning

The major types of knowledge we have distinguished are *normative knowledge*, *world knowledge*, *responsibility knowledge*, *reactive knowledge*, *creative knowledge* and *meta-legal knowledge*

Normative Knowledge. Normative knowledge is the most typical category of legal knowledge, to such an extent that to many authors 'normative' and 'legal' are practically the same thing. The basic conception of norm used in the ontology is largely derived from [11]. A norm expresses an idealization: what ought to be the case (or to happen), according to the will of the agent that created the norm. This view on norms is a generally accepted one, but it should be noted that this view does not make norms different from any abstraction. A model is also an idealization with respect to reality. What has been easily overlooked is that the idealization is motivated by desirability; not by a

match between model and reality. Norms limit possible behaviors and the assumption is that this subset of possible behaviors is a better world (from the perspective of politics). This implies also that this possible world is conceived in good correspondence with reality and that the norms refer to agent based activities that can be driven by intention, and not make wrong assumptions about physical and intentional processes.

Meta-Legal Knowledge. As stated above, we take the perspective that the law is defined on the basis of individual norms: e.g. individual articles in a regulation. The difference between the standard defined by the normative system on the one hand and the standards defined by the norms on the other hand is fundamental to understanding the role of normative knowledge in law, and it is accounted for by knowledge *about* individual norms. This distinction is captured by defining the categories of primary norms and meta-legal knowledge. Primary norms are entities that refer to agent caused, human behavior, and give it a normative status. This normative status is, in principle, either *allowed* (legal, desirable, permitted) or disallowed (illegal, undesirable, prohibited). However, each norm refers only to a few types of behavior, in the sense that it is only able to provide a status if applied to some types of situations (cases). For the remaining types of cases, the norm is said to be *silent*.

There may be a difference between the normative status given by a single norm and the one ultimately given by the normative system as a whole. Individual norms may conflict: if that is intended then some norm may be an exception to another norm. In order to solve these normative conflicts, meta-legal knowledge is applied. Typical conflict resolution is provided by meta-legal rules that state for instance that the more specific rule should be applied rather than a more general one: *"lex specialis derogat legi generali"* expresses this age old wisdom in law. Meta-legal knowledge is not only used for solving conflicts between norms. Another function is to specify which legal knowledge is *valid*. Validity is a concept which can be used both for specifying the dynamics of the legal system and its limits. A valid norm is one that belongs to the legal system (cf. [34]).

World Knowledge. By its very nature, law deals with agent caused behavior in the world. Therefore, it must contain some description of this behavior. For instance, in order to describe how the world should (ought to) be, norms must describe how things can be. In addition to adopting a category of legal knowledge which describes the world, we propose that this knowledge constitutes a structured model of some domain of law. For instance in a traffic act, the world of traffic is assumed to operate in a certain way, and the legislator may pose constraints – norms – on this behavior. A car may drive on all sides of a road, but the legislator (in most countries) obliges us to take the right hand side. Implicitly, the legislator has some model in mind of how traffic operates and can operate. The behavior in this world has to be modeled. This model is a generic one: how things (may) work, are done, or may be done in general if there are no normative limitations. In principle, the legislator has to foresee all possible types of situations and label these as allowed (desirable) or not. Thus, the term *legal abstract model* or LAM is used as a synonym for world knowledge when its model character is to be stressed.

The legal abstract model is an interface between the real world and the legal world. Its role is to define a model of the real world which is used as a basis to express normative

and other categories of legal knowledge. The bulk of the LAM consists of definitions of concepts that represent entities and relations in the world, i.e. it is to be viewed as an *ontology*.

Apart from describing the world, what is behavioral reasoning used for in law? We propose that this description of possible and relevant behaviors is built around the concept of *cause*, in order to allow the assignment of responsibility of an agent for a certain case. Causal knowledge, however, refers or uses a static description of the world (e.g. to model world states). Accordingly, we propose that the world model is actually composed of two related types of knowledge: terminological or *definitional knowledge* and *causal knowledge*. The definitional knowledge (ontology) is used by the normative knowledge to describe the ideal world they define. The causal knowledge is used by the responsibility knowledge to describe who or what have caused a given state of affairs, and can thus be considered responsible for it.

Responsibility Knowledge. Responsibility is the intermediary concept between normative and reactive knowledge, since a reaction can only occur if the agent is held responsible for a certain norm violation. Responsibility knowledge plays the role of linking causal connections with a responsibility connection — i.e. that connection which makes an agent accountable for a norm violation and possibly subject to legal reactions (see also Section 4.1). As [31] point out, however, responsibility does not have "any implication as to the type of factual connection between the person held responsible and the harm" — that is, causal connections are only a "non-tautologous ground or reason for saying that [an agent] is responsible" [31–pag. 66]. There are two basic mechanisms which are used in responsibility knowledge. First, the law may establish a responsibility connection independent of a causal connection — i.e. a *responsibility assignment*. This can be seen in a rule used in e.g. French, German or Brazilian law, by which parents are held responsible for the the damage done by their children even if there is no specific causal link between their attitudes or actions and the damage. That is, the parents are held responsible even though they have not necessarily caused the damage. Second, the law may limit the responsibility of an agent under certain circumstances, disregarding some possible causal connections — i.e. a *responsibility restriction*. For instance, in England a man is not guilty of murder if the victim dies more than one year after the attack, even if the death was a consequence of this attack. Other well-known factors that may influence the establishment of responsibility connections in law are *knowledge* and *intention*.

We refer here further to the work of [3] (see also [5] in this volume) who has worked out the relation between (physical and agent) causation and the various notions of responsibility of law; Lehmann has specified this in a foundational ontology that defines the notions of causality and causation.

Reactive Knowledge. To reach the conclusion that a certain situation is illegal (based on normative knowledge), and that there is some agent to blame for it (responsibility knowledge) would be probably useless if the legal system could not react toward this agent. That knowledge that specifies which reaction should be taken and how is what we call *reactive knowledge*. Usually this reaction is a sanction, but in some situations it may be a reward. The penal codes, which are usually a fundamental part of legal systems of

the Romano-Germanic tradition, contain basically responsibility and reactive knowledge only.

Creative Knowledge. A legislator may indirectly create some entity that did not exist before in the world, using what we call *creative knowledge*. It is usually stated in imperative terms, designating an agency that previously did not exist as part (or not) of the reality from a certain point of the time on. The creative function has a somewhat exceptional (or even abnormal) status if compared to the other ones. In this case, the law not only wants to classify or to react over certain agents that already exist in the real world, but attempts to create a new agent.

4.2 In Search of Ontological Foundations of Law

We have used the framework of *FOLaw* as a lead for fundamental research [35, 3], and as the basis for practical applications and architectures for legal reasoning (e.g., *ON-LINE* [4]). The CLIME project [6] was aimed at the construction of a legal information server. The try-out domain were international rules for safety and environmental care at sea, and the rules for ship classification (certification): in total about 15,000 different articles. This CLIME information server has two modes of operation. The first mode is typical information retrieval, where keywords (in phrases) are matched against terms in the rules. A large ontology (over 3,500 concepts) allows the elaboration of the keyword-terms by implied terms. The second, more expensive and experimental mode is in fact a question answering one. The CLIME system assesses whether a case, e.g. results of the inspection of a ship, or legal questions during the design of a ship, complies with the rules or not. The applicable (violated or 'potentially' violated) articles provide the justification and focus for the answer. An overview of CLIME and an evaluation of its results can be found in [36]. Other applications of the *FOLaw* framework (annex architecture) are reported in [37] (PROSA, a training system for solving legal cases); in the KDE project [7] the ontologies of CLIME have been re-used [38].

These results about the use and re-use of *FOLaw* also show its limitations. In developing the legal domain ontologies it turns out that the major effort in modeling is in the world knowledge. This is not surprising, given that the initial analysis about the content of legislation as in the TRACS project already revealed that world knowledge is the driving force in legal reasoning systems. More theoretically, it appears that law does not have its own ontological foundation. When legal philosophers discuss the ontological assumptions in law and legal reasoning, it is invariably about normative knowledge. This is different from other knowledge based fields of practice like medicine [39] or engineering [40], which have abstract ontological foundations in notions about physics, mathematics, etc. Jurisprudence and legal philosophy are primarily concerned with the *justification* of law and legal systems, rather than the *explanation* of the working of law and its relation to social reality. This is not to blame jurisprudence. The explanation of

[6] CLIME was an European project (IST 25414, 1998-2001): see http://www.bmtech.co.uk-/clime/index.html).

[7] KDE, for Knowledge worker Desktop Environment is a European IST project (IST 28678,1999-2001); see www.lri.jur.uva.nl/kde).

social reality is the concern of sociology, political and management science, etc. Explanations are models and these are grounded in ontological commitments. However, justification –which is derived from the term *ius* (law)– is the domain of epistemology; the study of what we can know and believe.

4.3 Epistemological Frameworks and Ontologies

Epistemology is about reasoning, argument and evidence, while ontology is concerned with modeling and explaining the world. Therefore, it is no surprise to see that 'core ontologies' about law are rather epistemic frameworks [41, 42]. [3] and [33], who make the same classical philosophical distinction between epistemology and ontology, construct 'ontologies' that mix both ontological and epistemological entities. In [43] we argue that this mixing up is theoretically not very clean, but the practical consequences may be a limitation of re-use and problems in interoperability. We want to maintain this classical distinction between ontology and epistemology. Epistemology is concerned with valid reasoning to arrive at justified conclusions. Of course, this reasoning is dependent on content: on the understanding (modeling) of the world, so the relationship is very intimate. However, the relationships between concepts is different from a perspective of ontology than from a perspective of epistemology. Typical epistemological terms, like hypothesis, evidence, conclusion, data can be reified as concepts in an ontology, but the epistemological relations are different. For instance, by using evidence on the basis of data a hypothesis may be confirmed and lead to a conclusion. So there are dependency relations between these reasoning states and these dependencies constitute argument structures or problem solving methods. However, this is not the way these terms are related in an ontology. Hypothesis and conclusion are roles in a problem solving method, while (empirical) evidence is a relation between states of a problem, as conveyed by these roles, and data. The backbone structure of an ontology consists of subsumption and mereological abstraction hierarchies, while epistemological structures are built from dependency and consistency relations between problem abstractions and data abstractions. Therefore *FOLaw* is rather to be viewed as an epistemological framework than as a core structure of legal ontologies. It can easily be re-written as a high level CommonKADS inference structure, as can be found for instance in [44]. [8]

5 Ontologies for Reuse: *LRI-Core*

If *FOLaw* is not sufficiently detailed and is rather an epistemological framework, there is a need for a new approach for constructing a core ontology for law. *FOLaw*, based upon notions of Kelsen, Hohfeld and Hart, has shown us two distinctive sets of concepts that are typical for law. (1) normative terms (and their definitions and axioms), and (2) responsibility terms (liability, guilt, causation, etc), which confirms what we have found in legal theory. An ontology of normative terms has been worked out by [1]. A foundational ontology that relates responsibility issues to agent– and to physical causation has

[8] It is curious to note that the first two authors of this article have worked on the design and content of this CommonKADS PSM library, but have not noticed this close formal correspondence.

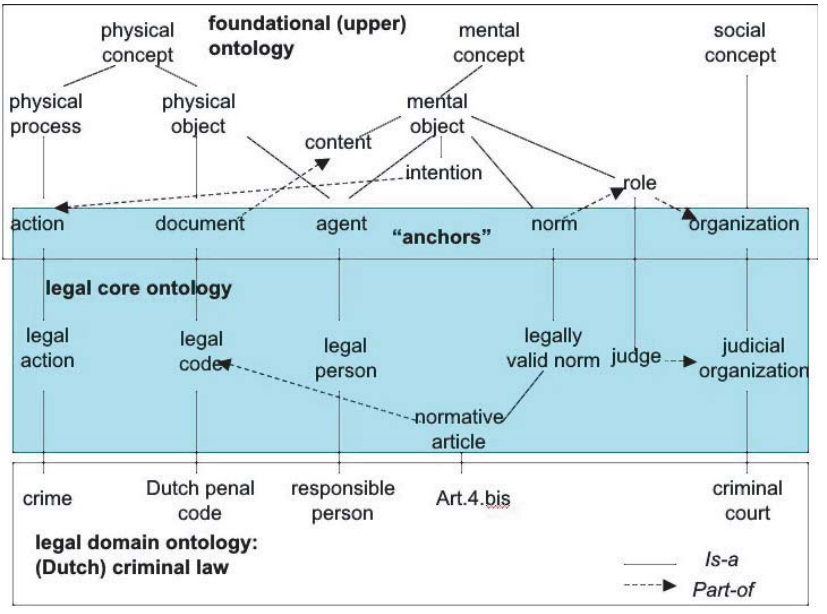

Fig. 2. LRI-Core layers: foundational and legal core share 'anchors' (high level concepts typical for law)

been constructed by [3]. Responsibility and norms may be notions that are typical for law, but they are not exclusive for law. The research of [45] (see also [46], this volume) shows that the terms of French law cannot be distinguished from common-sense terms, on the basis of statistical clustering methods. In other words, it appears that the law does not have a specific vocabulary, as is the case in other professional domains. If it is hard to find terms that are exclusive for law, we may be able to find terms that are still typical for law. The law is aimed at social activities. That means that notions of role, social position, and other social relationships and actions, in particular communicative ones, play a dominant role in domains of positive law. Property relationships, damage, individuality are some other recurring themes in law. Also law relies heavily on documentation and procedure. However, these typical, but not exclusive concepts in law are founded deeply in *common-sense*. That means that for modeling and understanding some legal domain we should be able to include notions about agents, actions, processes, time, space, etc, i.e. some foundational ontology appears to be indispensable on top of a core ontology of the *typical* legal terms, because the concepts of law are spread over almost the full range of common-sense. This means that a core ontology of law that covers these concepts that are typical for law should be grounded in some foundational ontology. Figure 2) sketches how *LRI-Core* is a legal specialization of some highly abstract common-sense concepts. In this section we will explain the perspective from which the common-sense grounding of *LRI-Core* is constructed, and present an overview of the main conceptualizations. In Section 6 we show how *LRI-Core* is applied to an ontology of (Dutch) criminal law.

We could not simply start with one of the currently available foundational or "upper" ontologies (e.g. Sowa's [47],or the IEEE-Standard Upper Ontology (SUO) [9] [48]) because their focus is rather on describing the physical and formal-mathematical world: not the social/communicative world which is more typical for law. These ontologies are 'revisionary' in the sense that they take ontology to represent our modern, scientific and formal insights as a point of departure, which is divergent from the 'naive' or 'folk' theories that make up common-sense. The upper part of the CYC [10] ontology and DOLCE [49, 50] are claimed to have a common-sense view, but this common-sense view is rather arbitrarily based upon personal intuition than on empirical evidence. For the construction of *LRI-Core* we have been inspired by results from studies in cognitive science: in particular evolutionary, neural and developmental psychology. Examples of how empirical evidence may support 'revisions' or 'folk' conceptualizations, are the following. In *LRI-Core* we will present in this section, space has two different meanings: 1) it refers to the *size* (extension) of physical objects, and 2) it refers to *positions* in space. This distinction is reflected by neural activation of different parts of the cortex [51]. On the other hand, contrary to common-sense wisdom, neurological evidence shows that we act before we decide to act in many circumstances, cf. [52]. This finding is so contrary to our common-sense intuitions that it would take a major overhaul of our folk psychology to make this part of our common-sense. However, common-sense is not an evolutionary given and fixed collection of conceptualizations: cultural revisions occur. For instance, although the notion of energy was introduced in physics just over a century ago, over the years it has become an undeniable part of our common sense conceptualizations. Another reason for not starting from already available foundational ontologies is that these do not very well cover the concepts typical for law such as roles, mental objects and processes, documents, social and communicative actions, etc. Extensions of DOLCE do, as one can see in the article by Gangemi *e*t al.} in this volume, but DOLCE was not available yet when we started development of *LRI-Core*. Moreover, like all other researchers in this area (and philosophers over the ages), we did not agree with the solutions proposed by these ontologies. The disagreement is not only due to a lack of grounding in evidence about common-sense distinctions. We also made a number of design choices which are different from other foundational ontologies. Some of these are:

- We do not make the distinction between 'perdurant' and 'endurant' entities as in Sowa's ontology and in DOLCE. In principle all concepts are endurants, i.e. all concepts are 'timeless'; all instances are perdurants (occurrences).
- Mental concepts are not "non-physical" concepts (DOLCE); the mental world is an analogon of the physical world with an intentional perspective.
- Energy is virtually absent in other ontologies. In *LRI-Core* it plays an important role in defining mental and physical processes.
- The notion of role covers in *LRI-Core* most social concepts, where in other ontologies role is rather a relationship (Sowa) or

[9] http://suo.ieee.org, A merged version of these ontologies, called SUMO, is available at http://ontology.teknowledge.com/

[10] www.cyc.com

The foundational layer of *LRI-Core* is not meant to be a fully worked out stand-alone foundational ontology. It contains relatively few concepts: we expect no more than about three-hundred concepts. Only those concepts that have a legal significance are fully worked out (the 'anchors' in Figure 2). The other concept are intended as a coherent coverage. *LRI-Core* is currently still under active development, it is expressed in OWL-DL, using the OWL-Plugin of Protégé [11]. A first version of *LRI-Core* has been developed and used to support knowledge acquisition in the e-Court project (see Section 6). In the next subsection we will present a short overview of *LRI-Core*.

5.1 Principles and Main Categories of *LRI-Core*

The top of *LRI-Core* consists at the moment of five major categories: each referring to a 'world'. These five are: physical and mental concepts, roles, abstract concepts and terms for occurrences. Most likely we have to add a sixth category: life (see below). These categories follow from an assumed evolution of human (and animal) conceptualizations of reality. Primary conceptualizations are inspired by moving and sensing, i.e. real-life interactions with the *physical* world. The complexity of this causal world is reduced when we take a 'teleological' stance with respect to life, in particular on living organisms of the same species. A teleological or intentional stance implies that the actions of agents are assumed to be motivated by goals. Teleological reasoning works 'backward', i.e. it allows reasoning from end-states (goals) to current states. This is less complex than the branching of possible worlds in causal, forward reasoning. Living creatures seek the maintenance and reproduction of life.

As human beings (and to some extent other higher mammals as well) discovered their own *mental* life, i.e. consciousness and self-awareness, the need arose for models of mental processes and objects. Awareness not only enables us to handle our own reasoning and emotions, but also to understand those of our fellow creatures in order to plan social activities and to communicate. Self-awareness enables 'reification', the building of metaphors that makes up *abstract* conceptualizations. These considerations convinced us that the mental world can be conceived as an intentional metaphor of the physical world, i.e. our mental life is made up of objects and processes. The categories we use to understand our own and other people's mental events mirror those of the physical world. The emergence of conscious planning and prediction of behavior has led to the conceptualization of *roles* that make up social organization. *LRI-Core* has thus been equipped with the following main categories: *physical, abstract* and *mental* concepts, and *roles* [12]. Finally, *LRI-Core* knows about a fifth category: *occurrences*. Strictly speaking, occurrences are not part of an ontology, as we will explain below. Figure 3 presents the top two layers of the ontology.

Occurrences. An ontology should not be structured according to the *way* things occur in physical, mental, or fantasy worlds, but rather to *what* the things 'essentially' are.

[11] See http://protege.stanford.edu

[12] One may argue that we have omitted another major category: life, or rather agent hood. Indeed, the distinction between non-living physical objects and living ones (agents) is crucial in common-sense. We have not (yet) investigated this category.

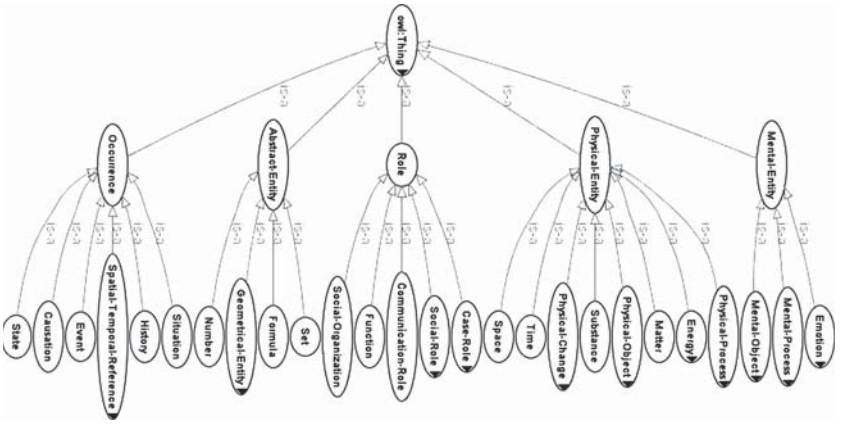

Fig. 3. LRI-Core, top two layers

Ontology has a Platonic flavor in the sense that it specifies the ideas with which we understand a/the world as it passes by. Making sense of the world means that we build models of current, past, and even to some extent, future situations. The structure of entities occurring in a world is different from the (abstraction) structure of generic concepts that make up an ontology. The concepts defined in an ontology enable us to recognize entities and their relations as they occur in the world. A simple example may illustrate this distinction. A functional ontology of furniture may distinguish as major categories objects one can sit on (chairs, banks); objects one can put objects on for immediate access (tables, desks), and objects for storing objects (cupboards, shelfs). So far so good, but the furniture in my room does not exhibit this structure at all. The structure to capture is in the first place a spatial one, i.e. dealing with positions. The distinction between these two views on furniture is not simply the distinction between instances and concepts, as one may imagine the generic spatial structure of a typical office versus a typical living room. Again, these structures, or rather frameworks or models, are different from those in an ontology of furniture. Note that we have met already another kind of framework: the epistemological ones, which we wanted to distinguish from ontologies in Section 4.3. In principle, as argued above, in a clean ontology there is no place for frameworks or generic models. Both can be part of a knowledge base, but for purposes of clarity and re-use one would like to keep them apart.

Ontologies define and deliver the building blocks for the construction of interpretations of actual situations and histories: partial models of real or imaginary worlds. Histories describe the life line of individual entities, and situations are diachronic spatial structures of objects and processes. The distinction between situation models and the concepts we use to identify the elements (parts) of situations, is obscured in ontologies that make a fundamental distinction between occurrents (perdurants) and continuants (endurants) (e.g. Sowa, DOLCE). Perdurants are entities that have parts that change with time or place. “For example, the first movement of the (execution of) a symphony is a temporal part of it.” ([53–p. 20]. Indeed the execution of a symphony has temporal parts, but the concept of symphony itself has not. One may hold that all entities (instances, individuals) that exist are perdurants. Even a stone, a typical ‘endurant’ in these ontolo-

gies is in the execution of its life-line also perduring. Originally part of a rock, a stone may end up as sand on a beach, gradually spreading its parts spatially. However, when we take the concept of first-movement-of-a-symphony, or (pars-pro-toto) a symphony, there is nothing of temporal parts, even if the score represents a temporal sequence. The notion of symphony, once created, remains an 'eternal' concept. It does not differ from our notions of a process or a physical object. Strictly speaking, all entities in situations are endurants; all concepts are perdurants.

The distinction between a concept and its occurrence is particularly relevant where mental concepts (here: entities) get executed. For instance, plans, norms and roles and their execution (respectively, their observation) may show divergences that can be marked as 'bad' or even illegal. A divergence can only be identified if a mental plan, norm or role still exists so that it can be compared with actual behavior (or its memory or recording).

The category of occurrences in *LRI-Core* captures those strictly temporal aspects related to the *execution* of scenarios involving objects and processes. This means that events are occurrences, but processes are not. Where processes contain the explanation of the changes they cause, events only describe a discrete difference between the situation before and after the event took place: they describe the input-output of the execution of a process, and happen 'in' time. All this does not reduce the need for terms to talk about occurrences in general. For instance, above we have used terms like situation, event, history and entity. These terms refer to occurrences in an abstract sense that can legitimately be part of an ontology that defines concepts. Therefore, LRI has a category of 'occurrences'.

The distinction between real occurrences and ontology is reflected in a major distinction in human memory. Psychological research has identified two types of 'declarative' memories. The distinction between semantic memory and episodic memory is well established (see e.g. [54]). Semantic memory corresponds with our knowledge about the world, i.e. ontology; episodic memory contains memories of events (occurrences). Semantic memory emerges in child development earlier than episodic memory: we have to know before we can understand events and situations.

Physical Entities. The physical world evolves around two main classes: physical objects and processes. Objects are bits of matter, which in turn is typed by its substance. Objects have mass, extension, viz. form and aggregation state (limiting form). The existence of objects expresses the notion that matter (in particular solid matter) is what renders the physical world relatively stable and observable. Physical situations are usually described by the arrangement of instances of physical objects.

This intuition does not exist for the second major class that governs the physical world: process. Processes consume energy to change objects, or parts of objects. Though energy is a naively problematic concept (See [55]), its use has become widespread to such an extent, that it has conquered its place in common sense. Particularly the fact that electricity can be converted in many types of energy has enabled the common acceptance of this 'revisionary' concept. Processes are described by the changes they bring about. Change is an inherently temporal concept, belonging to the realm of occurrences. Through interaction, processes can cause one another, leading to series of events that only stop at some equilibrium: in general conceived as that there are no interactions at all. In *LRI-Core*, processes are typed according to two views: (1) formal change (transfor-

mation, transduction and transfer) and (2) the kinds of (properties of) objects involved. (e.g. movements are the change of position of objects; chemical processes change the substance of objects, etc.). A third property is whether a process produces or consumes energy; the default is the latter.

The concept of process is often used as synonymous to action and activity. LRI-Core defines actions as processes that are initiated by an agent acting as actor. Notwithstanding the intricacies of mental (or agent) causation ([56]), the action itself is strictly physical: i.e. some muscle movement. The mental perspective implied by agent-causation is that actions are intended: they are preceded by some kind of intentional decision to act. Many ontologies use the term process to cover both processes and actions. Business processes, to give an example, consist of actions. By abstracting out the agents one may see the work in an organization as anonymous processes, but they do not exhibit the same causal gluing as in physical processes. For instance, business processes are planned and controlled, i.e. initiated by (supervising) agents. The analogy for planned activities is the causal design of a device. Devices funnel causal chains of processes in such a way that they exhibit intended behavior, expressed as the functions of a device.

Mental Entities. Conceptions of the mental world have a strong analogy to the physical world. We conceive the mind as consisting of (mental) objects, like concepts and memories, which are processed by mental processes that transform or transfer these objects. Memories are retrieved; concepts are formed. Moreover, these mental objects may be aggregations of more elementary objects. Memories consists of multi-media representations of situations experienced; thoughts are made of more elementary parts like concepts.

The contents (substance) of these objects are representations. The conceptual content of thoughts is intended by propositional attitudes, like belief, desire, norm etc. Mental objects are processed or stored in containers (such as the mind) which in turn can have parts, e.g. memories. Mental processes like thinking, memorizing, imaging are operations on mental objects. The equivalent of physical energy in mental processing is the concept of emotion: the forces that make us focus our mental energies.

There is however, an important difference between the mental and the physical. Where physical processes are governed by causation, mental processes are controlled by intention. If that is the case, we would rather use the term ‘action’ for these processes. Thinking is thought to be an action, as we assume that we have full control over our thoughts and can decide about what we are thinking. However, where our mind escapes our conscious intentions, as e.g. in getting in uncontrollable emotional states, or in forgetting an appointment, we rather take a physical than an intentional stance. Despite this subtle difference, we keep the term mental process to cover both, as we want to reserve the term action for agent-caused processes.

The outcome of a mental process can be the intention to act, for instance according to a structure of primary actions: a plan. These actions can be aimed at bringing about both physical and mental changes, e.g. changing the mental state of another agent. Such intended mentalistic actions are acts of communication (which also need some physical medium to transfer the intended mental state). Speech act is the most common of these actions.

The role of mental conceptualizations is extremely important in understanding and communicating with other people. Their primary use lies in their role as building stones of models of the minds of other people: user-models. The intentional stance means that we attribute intentions and intention directed mental processing and belief to other people and to some extent animals (or even computers, [57]).

Roles. Roles cover functional views on physical objects (devices), on agent behavior or on mental processes. In particular, social behavior and social organizations are explained as (consisting of) roles. Typical mental roles are epistemological ones. For instance conclusion, evidence and hypothesis are roles in problem solving processes and can therefore also be categorized under mental classes [58]. From a role perspective, functions are roles of physical objects, e.g. we may use objects for non-intended functions.

Roles are entities in the mind, they do not 'really' exist. Roles are idealizations: we may not play a role correctly. An important distinction should be made between playing a role and the role itself: "agents can act, and roles cannot" [59]. Correcting incorrect role playing does not mean that we change the role: we change our behavior. Like plans and processes, roles in ontologies are often confounded with their execution, in the same way as the execution of a symphony may be confounded with the symphony itself. The original meaning of the term role refers to a role of paper that contained the text of an actor in a play. Also the role-taker (some agent) and the role are often confounded, which may become obvious when we identify a role with a person. These kinds of confusions have made conceptual modelers aware of the tricky issues about roles (see e.g.[60]).

Roles are often viewed as relationships ([47, 60, 61]). Indeed, social roles have mutuality and complementarity. No students without teachers; no parents without children; no speakers without hearers, etc. In theory of law, a related view exists about the mutuality of legal positions: i.e. rights and duties [22, 11]. For instance, if citizens have the obligation to vote, the government has the duty to enable this voting. Nevertheless, this complementarity of roles might not be of enough importance to grant their representation as relationships in an ontology. The ontology may specify such relationships, but the primary notion of role is as a concept.

This becomes clear when we look at roles as concepts, i.e. at what roles *mean*. Roles are behavioral requirements on role execution and on qualifications of role taking. These requirements are prescriptions, i.e. they are normative. In modern society many roles have formal requirements enforced by *law*. Legislation addresses actors by the roles they play [13]. If actual behavior deviates from the norms attached to these roles we violate the law. Violations are based upon the distinction between the prescription (role) and role performance. Therefore, in court, it is the actor of the role who is made responsible: as a person; not as a role. Even the fictitious concept of legal-person for social organizations turns into concrete responsibilities of the liable persons who have mis-performed their roles.

Abstract Entities. As all concepts are abstractions, one may argue that a separate abstract world is difficult to see. However, common sense knows about a (small) number of proto-mathematical concepts, such as collections, sequences and count-numbers (positive integers). We know about geometric simplifications such as line, circle, square,

[13] An exception to this rule is in criminal law.

cube, etc. [62] even argue that these common sense notions are the real roots of our mathematics. Nonetheless, these kind of semi-formal abstractions do not play a very central role in law, and therefore *LRI-Core* is thinly populated with abstract classes.

The role of the concepts defined in *LRI-Core* is illustrated in the next section, where we present fragments of an ontology of Dutch criminal law, as developed in the e-Court project.

6 An Ontology of Criminal Law: *OCL.NL*

The e-COURT project was aimed at the semi-automated information management of documents produced during a criminal trial: in particular the transcriptions of hearings. The structure of this type of document is determined by the debate/dialog nature of these hearings, but also by specific, local court procedures. Besides tagging its structure, it is also important to identify (annotate) content topics of a document. These vary from case descriptions (e.g., in oral testifying) to topics from criminal law (e.g., in the indictment).

The case descriptions have a strong common-sense flavor, but the legal professionals who are the main intended users are primarily interested in the (criminal) legal aspects of a case. We developed an ontology that covers Dutch criminal law, whose major structure we will discuss below. As the e-COURT solutions are aimed to work for most European countries, in principle one has to develop such an ontology for every jurisdiction that intends to use e-COURT. This Dutch ontology was intended to work as a reference for the development of similar ontologies of Italian and of Polish criminal law.

We can illustrate the use of 'anchors' in the *LRI-Core* ontology with parts of the ontology for Dutch criminal law (OCL.NL). In Figure 4 the boldface terms are terms from *LRI-Core*. *LRI-Core* knows about the distinction between a person as a lifetime identity and roles that a person may perform. Roles are taken by persons who are agents. Agents are both physical and mental objects: dependent on the context of use the physical or the mental properties of an agent are selected. This solution is more elegant than assuming that an agent has a body and a mind. That is the Cartesian solution. The two views on 'agent' correspond better with a more unified view of the identity of an agent [63, 64].

In Figure 5 a selection of typical legal roles in criminal (procedural) law is presented. In *LRI-Core* we distinguish between social roles and social functions. Social functions

agent
| **person**
| | natural person
| | juristic-person
| | | company
| | | association
| | | foundation
| **collection-of-agents**
| | **group**

Fig. 4. Agents in Dutch Criminal Law (OCL.NL) (excerpt)

role
- **social-role**
  - legal-role
    - juridical-role
      - judicial-role
        - judge
          - judge-presiding
        - prosecution-role
          - public-prosecutor
        - defense-role
          - defense-counselor
        - defendant
          - principal-defendant
          - accessory-defendant
          - offender
            - convict
        - witness
      - clerk-of-court
      - lawyer
      - the-Regent
      - the-State
    - public-servant
    - owner-of-goods
    - owner-of-rights/duties
      - creditor
      - debtor
- **social-organization**
  - legal-organization
    - public
      - Ministry-of-Justice
      - courts-by-jurisdiction
        - criminal-court
        - administrative-court
      - courts-by-level
        - cantonal-court
        - court-of-appeal
        - Supreme-court
- **social-function**
  - public-social-function
    - jurisdiction
      - public-prosecution
    - criminal-investigation
      - forensic-investigation

Fig. 5. roles and functions in Dutch Criminal Law (OCL.NL) (excerpt)

are external roles of organizations. Social roles make up the functional internal structure of an organization. In these figures we cannot show multiple classification, nor other relations between classes than subsumption. For instance, an organization has social

mental-object
| juridical-mental-object
| | legal-norm
| | judicial-mental-object
| | complaint
| | accusation
| | judicial-decision
| | | verdict
| | | | conviction
| | | | acquit
| | | | final-verdict
| | juridical-qualification
| | | deontic-qualification
| | | | deontic-legal-role-attribute
| | | | | right
| | | | | duty
| | | | | authority
| | | | deontic-modalities-of-norms
| | | | | permission
| | | | | obligation
| | | | | prohibition
| **reasoning-role**
| | **evidence**
| | | testimony
| | | | eye-witness-testimony
| | | forensic-evidence
| | **problem-solving-role**
| | | **solution**
| | | **problem**
| | | **problem-solving-method**
| | **argumentation-roles**
| | | **debate-argument-role**
| | | | accusation-position
| | | | defense-position
mental-process/action
| **internal-mental-processes**
| | **reasoning**
| | **communicative-mental-action**
| | testifying
| | interrogating
| | **argument**
| | **dialog**
| | | **dialogical argument**
| | | | dispute
| | | | | judicial-dispute

Fig. 6. Mental objects, processes and states in Dutch Criminal Law (OCL.NL) (excerpt)

functions and 'has-as-parts' social roles. This is not the only view on the composition of an organization. The hierarchy of authority is another one, but this hierarchy maps onto the roles: authority is a mental entity: to be precise a 'deontic-legal-role-attribute' Figure 6 presents some of the major categories of the mental world.

In this representation of the mental world we have skipped some views. Some mental legal objects, such as 'accusation' are in fact (illocutionary) acts. In legal discourse an accusation is really treated as an object, i.e. it is the (content; sometimes the literal surface structure) text that is referred to. However, its meaning is indeed the act of accusation, so it should inherit properties of mental objects and those of (illocutionary) mental actions. Legal procedures may objectify or 'reify' these actions.

Many objects of the mental world are reifications of epistemological roles. Terms like 'reason', 'evidence','explanation', 'problem', 'dispute' etc. come from the vocabulary of reasoning methods and are concerned with assessing the (trust in) the truth of (new) beliefs. As stated in the Introduction, law is particularly with justifying legal decisions, so roles in argumentation and reasoning play an important role in legal discourse in general. Note that these terms are not part of Dutch criminal law, except some global terms like 'evidence' and 'doubt'. We have added some because they occur frequently in court discourse.

The hard core of the OCL.NL consists of actions. There are two major types: the criminal actions themselves (called 'offenses'). These are of course the actions executed by the person who is successively acting as suspect, defendant, and eventually convict (if true and proved. . .). On the other side, the convict may be at the receiving end of the 'punishment' actions, that are declared by the legal system etc. Crime and punishment are the keys to criminal law that is synonym to penal law.

6.1 Use of Ontologies in Legal Information Retrieval

The ontologies of criminal laws are used in e-COURT to support the information retrieval of information contained in the hearing session documents. Criminal law is only part of the discourse in these sessions, but an important part. Another part consists of descriptions of 'what has happened' for which only full blown common sense (CYC) or superficial but extensive ontologies like Wordnet may play a role in information retrieval. Thus far we focus on the criminal legal terms, because the primary type of users are legal professionals.

In e-Court, two user modes of search are used: basic and advanced. The basic search mode allows meta-data and/or keyword search by specifying values for one or more meta-data fields and/or keywords. The advanced search mode includes possibilities to use linguistic weights and quantifiers with the keywords, to select the language of the query and the searched documents; to choose particular document sections of interest. In this subsection we describe the specific additional information management functions that are supported by ontologies.

Annotation and XML Tagging of Legal Documents. In information management the emphasis has been on archiving and retrieving documents by their formal, syntactic characteristics. These structures are abstracted in meta-data: RDBM schemas, DTDs for XML-tags, XML-Schemata, etc. This works fine as long as the structures are rather

fixed and the occurrence of parts ('sections') is easy to identify in an automatic way. The criminal trial *hearing* documents in e-COURT are not the typical kind of documents that are handled by information systems. Hearing documents reflect in the first place oral, often 'spontaneous' **dialog** from the court room. The role of ontologies in indexing the e-Court hearing documents is threefold:

- The first role is an indirect one: the ontologies provide the structured vocabulary for meta-data descriptions and maintain consistent use and semantic distinctions. The XML-Schemata only provide 'syntactic', structural information, but the ontologies (expressed in OWL/RDFS) enable semantic coherence and verification.
- The identification of dialog-turns can be (almost) fully automated by the use of simple voice-recognition devices that have only to distinguish voice characteristics of the participants in the dialog. However, all other tagging of documents has to be done by hand by the transcribers of the hearing recordings. An ontology browser supports this activity.
- The e-COURT system indexes all documents. A number of these indexed terms correspond with terms of the ontologies. In this way we can link documents automatically with some semantics, i.e. one may gather what the document is about, which is functionally equivalent to (XML)-tagging the document with these

Query Expansion. The set of keywords used in a query can yield unsatisfactory results because the actual use of terms in a document may not correspond to what the user has in mind. This is obvious in the use of synonyms. However, also more abstract terms may be used to denote a more specific object: e.g. *killing* (synonym: *manslaughter*) for *murder*. A reference to a *murder* may be missed because in the document the terms *killing* and *manslaughter* are used. The reverse may also be relevant in information retrieval. The user may search for the *weapon* that is used in a particular criminal case, but may not know what kind of weapon exactly was used. By browsing a taxonomy of weapons (e.g. as part of an ontology of terms in criminal law) she may specify the query further. We have observed that users of legal information retrieval systems have a tendency to *under-specify* the cases they present. They do not provide all potentially relevant facts, and they use terms that are too general. Therefore the system may miss potential exceptions in the set of norms, and the user may deduce a wrong normative assessment of her case. Having a user interface that explicitly allows for specification of used terms may help. Another solution is to have the system return potential exceptions as well as norms that exactly match the case at hand ([34]). In both search modes (basic and advanced) the ontology repository is consulted for subsumed or subsuming terms with respect to the keywords given.

Expansion by Subsuming Classes By adding terms for searching that are superclasses of the already specified terms, the search is directed also to the more general, abstract terms. In searching documents that contain regulations (laws, statutes, contracts) where applicable provisions are often formulated in generalized and abstract terms, this IR strategy is in fact the only one to avoid false negatives (i.e. missed applicable provisions). In the CLIME project this strategy has been implemented [36].

Expansion by Subsumed Classes The example of the search for a *weapon* above shows the problem when the user is searching for a subclass of a term she may well know. There

are two possibilities. The user may allow all subsumed terms to participate as keywords in the search (which may lead to an explosive return of candidates) or she may have already restricted the set of possible documents and have a look at those *weapons* that occur as indices of these documents. In fact, the example is typical for the kind of searches where one is looking for additional, very specific information that should answer a question.

Disambiguation of a keyword term is another role of ontologies in information retrieval. Classical ambiguity consists of terms that have different meanings but the same 'orthography'. Except for orthographic coincidences, most ambiguous terms in fact share meaning, besides their differences. Disambiguation occurs in the context of use and is a matter of 'degree'. There may be little ambiguity in the term *car* as an isolated term, but there is little overlap in what it implies between the mechanic's and the salesman's view of *cars*, even if they work for the same company. In ontologies persistent, but context (role) dependent ambiguity is represented as **multiple classification**.

Except for disambiguation and selective use of terms of subsumed classes, the additional terms are added as disjunctive keywords to the query set, which means that the set of documents that is returned – the *'result set'* – may have increased exponentially. One may find more correct returns, but one must be prepared for a large amount of false positives: the classical problem of information overload we try to avoid and for which the major web stakeholders (at least the W3C) see the solution in the semantic web technology. It appears there is not a free lunch at the web, nor at e-COURT that seeks the same solutions. There are two methods to cope with this problem. The first one is to have the user refine his query. However, this is often a problem because the user may not have enough information to do this.

A second solution consists of *(re)organizing the result set*. The typical problem in (WWW) information search is that the number of returned documents may be unmanageably large and heterogeneous. The cause of much heterogeneity is the fact that a term may have multiple senses/views. In particular, the legal (criminal) domain is full of multiple views as we explained in Section 6, so we expect that disambiguation may occur by not only matching the indices of the returned documents with the keywords, but also have a second filtering/clustering where we also match indices with associated terms in the ontologies, i.e. the *value(-class)* and other related terms in the ontologies.

7 Discussion and Conclusions

The research reported here covers a long period – about fifteen years – of a number of mid-size (European) projects. In this article we have only pursued and discussed the views and results related to the development of a core ontology that identifies the main concepts that are typical for law. This guided tour has revealed that in fact the law has only a few of those concepts. Legal theory, but also work in AI & Law indicate that the most typical ones are concerned with normative knowledge (deontic terms) and with notions about legal responsibility [3]. Already at the start of our investigations in the TRACS project that was about traffic regulations we found that by far the majority of terms referred to the common-sense world of traffic; the only exception were the already mentioned deontic terms. Does this mean that law is a typical common-sense domain?

The answer is 'yes' and 'no'. The 'no' is explained by the fact that the legal world knowledge (LAM, in terms of Valente's *FOLaw* [1]) is a filtered and adapted version of what may have started as a common-sense. It should be noted that we have not divided up law in the way it is universally conceived and taught at law school, where the first distinction is between private and public law, and public law covers such legal domains as administrative law and criminal law. In each of these domains of law one will find concepts that may have evolved from common-sense, but which have received a typical and exclusive meaning in law. Moreover, as in all domains of professional practice, new concepts may have been developed. The understanding and use of the current state of these legal-domain specific concepts is for instance what legal education is about. A major reason for the evolution of legal terms in specific domains of law is in the first place due to the fact that the power of law is limited: it cannot command the physical world (so what is desirable in law is always a subset of what is possible). Therefore, only those concepts and relations are object of law that can be affected by human conscious (and individual) intervention. A second reason is that common sense terms may get refined and redefined in such a way that they correspond better with principles of law. A third reason is that law transpires the views and goals of the recent and current politics.

Therefore it appears that within legal domains much common-sense is filtered, 'cleaned' and transformed into a layer only understandable and usable for legal professionals. However, this is only part of the story. The legal system needs a close correspondence between the common sense view of the world and its terminology. This is not only required for drafting legislation, but particularly for citizens to understand the law and for interpreting cases. Cases are accounts of what has happened and they are cast in narrative discourse: events that are connected by causes and intentions (reasons). Therefore, a mapping between legal and common-sense concepts has to be maintained. All this is to say that a core ontology for law in general has problems in covering this large area of world knowledge and has to resort in a first approach to common-sense foundational ontologies. We have also identified terms which are not exclusive, but still very typical and well elaborated for law: document, document-structure, role, etc. are central terms in law and may be grounded in a core ontology that imports these notions from a still high level and simple foundational (common-sense) ontology.

References

[1] Valente, A.: Legal knowledge engineering: A modelling approach. IOS Press, Amsterdam, The Netherlands (1995)
[2] den Haan, N.: Automated Legal Reasoning. PhD thesis, University of Amsterdam (1996)
[3] Lehmann, J.: Causation in Artificial Intelligence and Law: a modelling approach. PhD thesis, University of Amsterdam (2003)
[4] Valente, A., Breuker, J., Brouwer, P.: Legal modelling and automated reasoning with ON-LINE. International Journal of Human Computer Studies **51** (1999) 1079–1126
[5] Lehmann, J., Breuker, J., Brouwer, P.: CAUSATIONT: causation in AI & Law. In: this volume. (2004)
[6] Breuker, J., Den Haan, N.: Separating world and regulation knowledge: where is the logic? In: Proceedings of the third international conference on AI and Law, Oxford, ACM (1991)
[7] Clancey, W.: Heuristic classification. Artificial Intelligence **27** (1985) 289–350

[8] Patil, R.: Artificial intelligence techniques for diagnostic reasoning in medicine. In Shobe, H., AAAI, eds.: Exploring Artificial Intelligence: Survey Talks from the National Conferences on Artificial Intelligence. Morgan Kaufmann, San Mateo, California (1988) 347–379
[9] Rector, A.: Medical informatics. In Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P., eds.: Description Logic Handbook. Cambridge University Press (2002) 415–435 http://www.cs.man.ac.uk/ franconi/dl/course/dlhb/dlhb-10.pdf.
[10] Raz, J.: Legal principles and the limits of law. The Yale Law Journal **81** (1972) 823–854
[11] Kelsen, H.: General Theory of Norms. Clarendon Press, Oxford (1991)
[12] Boer, A., van Engers, T.: A knowledge engineering approach to comparing legislation. In Wimmer, M., ed.: Proceedings of the Conference on Knowledge Management in E-Government, Springer (2003)
[13] Van Der Velden, W.: Coherence in law: a semantic and deductive explanation. In Brouwer, B., Hol, T., Soeteman, A., Van Der Velden, W., De Wild, A., eds.: Coherence and Conflict in Law, Deventer, Boston / Zwolle, Kluwer Law and Taxation Publishers / W.E.J. Tjeenk Willink (1992) 257–287 Proceedings of the Third Benelux-Scandinavian Symposium in Legal Theory, Amsterdam, January 3-5, 1991.
[14] Winkels, R., Haan, N.D.: Automated legislative drafting: Generating paraphrases of legislation. In: Proceedings of the Fifth International Conference on AI and Law, College Park, Maryland, ACM (1995) 112–118
[15] Den Haan, N., Winkels, R.: The Deep Structure of Law. In Prakken, H., Muntjewerff, A., Soeteman, A., eds.: Legal Knowledge Based Systems: The Relation with Legal Theory – Seventh International Conference on Legal Knowledge-based Systems Legal Knowledge-Based Systems, JURIX–1994, Koninklijke Vermande (1994) 43–54
[16] Breuker, J., Petkov, E., Winkels, R.: Drafting and validating regulations: the inevitable use of intelligent tools. In Cerri, S., Dochev, D., eds.: Artificial Intelligence: methodology, systems and applications. Springer (2000) 21–34
[17] Lindahl, L.: Conflicts in Systems of Legal Norms: A Logical Point of View. In Brouwer, B., Hol, T., Soeteman, A., Velden, W.V.D., Wild, A.D., eds.: Coherence and Conflict in Law, Deventer, Boston / Zwolle, Kluwer Law and Taxation Publishers / W.E.J. Tjeenk Willink (1992) 39–64 Proceedings of the Third Benelux-Scandinavian Symposium in Legal Theory, Amsterdam, January 3-5, 1991.
[18] Hart, H.: The Concept of Law. Clarendon Press, Oxford (1961)
[19] Bentham, J.: An Introduction to the Principles of Morals and Legislation. Athlone Press, London (1970) Edited by J. Burns and H. Hart.
[20] Austin, J.: The Province of Jurisprudence Determined. Weidenfeld and Nicolson, London (1954) edited by H.L.A. Hart.
[21] Hofstadter, A., McKinsey, J.: On the logic of imperatives. Philosophy of Science **6** (1939) 446–457
[22] Hohfeld, W.: Fundamental Legal Conceptions as Applied in Legal Reasoning. Yale University Press (1919) Edited by W.W. Cook, fourth printing, 1966.
[23] Allen, L., Saxon, C.: More IA needed in AI: Interpretation assistance for coping with the problem of multiple structural interpretations. In: Proc. of the Third International Conference on AI and Law, Oxford (1991) 53–61
[24] Allen, L.: The language of legal relations, RLL: useful in a legal ontology toolkit? In Visser, P., Winkels, R., eds.: Legal Ontologies: proceedings of the ICAIL-97 workshop. ACM, New York (1997) 47 – 60
[25] Hamfelt, A., Barklund, J.: Metaprogramming for representation of legal principles. In: Proceedings of Meta'90. (1990) 105–122
[26] McCarty, T.: A language for legal discourse I. basic structures. In: Proc. of the Second International Conference on AI and Law, Vancouver, Acm (1989) 180–189

[27] Alchourrón, C., Bulygin, E.: Normative Systems. Springer-Verlag, Wien (1971)
[28] Alchourrón, C., Bulygin, E.: The expressive conception of norms. In Hilpinen, R., ed.: New Studies in Deontic Logic. D. Reidel (1981) 95–124
[29] Kralingen, R.W.V.: Frame-based Conceptual Models of Statute Law. PhD thesis, University of Leiden, The Hague, The Netherlands (1995)
[30] Brouwer, P.: Samenhang in Recht: een analytische studie (coherence in law). Wolters-Noordhoff (1990)
[31] Hart, H., Honoré, T.: Causation in the Law. Second edition edn. Oxford University Press, New York (1985)
[32] McNamara, P., Prakken, H., eds.: Norms, logics and information systems; new studies in deontic logic and computer science, IOS-Press, Amsterdam (1999)
[33] Mommers, L.: Applied legal epistemology: building a knowledge-based ontology of the legal domain. PhD thesis, University of Leiden (2002)
[34] Winkels, R., Bosscher, D., Boer, A., Breuker, J.: Generating exception structures for legal information serving. In Gordon, T., ed.: Proceedings of the Seventh International Conference on Artificial Intelligence and Law (ICAIL-99), ACM, New York (1999) 182–189
[35] Lehmann, J., Breuker, J.: On defining ontologies and typologies of objects and processes for causal reasoning. In Pease, A., Menzel, C., Uschold, M., Obrst, L., eds.: Proceedings of IEEE Standard Upper Ontology, Menlo Park, AAAI-Press (2001) 31 – 36
[36] Winkels, R., Boer, A., Hoekstra, R.: CLIME: lessons learned in legal information serving. In Harmelen, F.V., ed.: Proceedings of the European Conference on Artificial Intelligence-2002, Lyon (F), Amsterdam, IOS-Press (2002)
[37] Muntjewerff, A., Breuker, J.A.: Evaluating PROSA, a system to train legal cases. In Moore, J., Redfield, C., Johnson, L., eds.: Artificial Intelligence in Education. FAIA-Series, Amsterdam, IOS-Press (2001) 278–290 ISSN:0922-6389, ISBN:1 58603 173 2.
[38] Jansweijer, W., Van Der Stadt, E., Van Lieshout, J., Breuker, J.: Knowledgeable information brokering. In Stanford, B., Kidd, P., eds.: E-bussiness;: key issues, applications and technologies. IOS-Press (2000)
[39] van Heijst, G., Schreiber, A.T., Wielinga, B.: Using explicit ontologies for kbs development. International Journal of Human-Computer Studies **46** (1997) 183–292
[40] Borst, P., Akkermans, J.M., Top, J.L.: Engineering ontologies. International Journal of Human Computer Studies **46** (1997) 365–406
[41] Van Kralingen, R., Visser, P., Bench-Capon, T., H, V.: A principled approach to developing legal knowledge systems. International Journal of Human Computer Studies **51** (1999) 1127–1154
[42] Hage, J., Verheij, B.: The law as a dynamic interconnected system of states of affairs: a legal top ontology. International Journal of Human Computer Studies **51** (1999) 1034–1077
[43] Breuker, J., Hoekstra, R.: Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law. In: Proceedings of EKAW Workshop on Core ontologies, http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/, CEUR (2004)
[44] Breuker, J., Van De Velde, W., eds.: CommonKADS Library for Expertise Modeling: reusable problem solving components. IOS-Press/Ohmsha, Amsterdam/Tokyo (1994)
[45] Lame, G.: Construction dÕontologie à partir de textes. Une ontologie de droit dédié à la recherche dÕinformation sur le Web. PhD thesis, Ecole des Mines, Paris, Http://www.cri.ensmp.fr/ (2002)
[46] Lame, G.: Using NLP techniques to identify legal ontology components: Concepts and relations. In: this volume. (2004)
[47] Sowa, J.F.: Knowledge Representation: Logical Philosophical, and Computational Foundations. Brooks Cole Publishing Co, Pacific Grove, CA (2000)
[48] Pease, A., Niles, I.: IEEE Standard Upper Ontology: A progress report. Knowledge Engineering Review **17** (2002) 65–70 Special Issue on Ontologies and Agents.

[49] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A.: Sweetening WORDNET with DOLCE. AI Magazine **24** (2003) 13–24
[50] Massolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., Schneider, L.: The WonderWeb foundational ontologies: preliminary report. Technical Report Deliverable D17, version 2, ISTC-CNR (Italy) (2002)
[51] Pinker, S.: How the mind works. Penguin Books (1997)
[52] Wegner, D.: The Illusion of Conscious Will. Bradford Books, MIT Press, Cambridge, Massachusetts/London, England (2002)
[53] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening ontologies with DOLCE. In Gomez-Perez, A., Benjamins, V., eds.: Proceedings of the EKAW-2002, Springer (2002) 166–181
[54] Tulving, E.: Concepts of human memory. In L.Squire, Lynx, G., Weinberger, N., McGaugh, J., eds.: Memory: organization and locus of change. Oxford University Press (1991) 3–32
[55] Hayes, P.J.: The second naive physics manifesto. In Hobbs, J.R., Moore, R.C., eds.: Formal Theories of the Common Sense World. Ablex Publishing Corporation, Norwood (1985) 1–36
[56] Kim, J.: Mind in a physical world: An essay on the mind-body problem and mental causation. In: Representation and Mind. MIT Press (1998)
[57] Dennett, D.: The Intentional Stance. MIT-Press (1987)
[58] Breuker, J.: Components of problem solving. In Steels, L., Schreiber, G., van de Velde, W., eds.: A Future for Knowledge Acquisition: proceedings of the EKAW-94, European Knowledge Acquisition Workshop, Berlin, Springer Verlag (1994) 118 – 136
[59] Pacheco, O., Carmo, J.: A role based model for the normative specification of organized collective agency and agents interaction. Journal of Autonomous Agents and Multi-Agent Systems **6** (2003) 145–184
[60] Steimann, F.: On the representation of roles in object-oriented and conceptual modelling. Data and Knowledge Engineering **35** (2000) 83–106
[61] Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N.: Social roles and their descriptions. In: Proceedings of Knowledge Representation Workshop. (2004)
[62] Lakoff, G., Núñez, R.: Where Mathematics Comes From. Basic Books (2000)
[63] Strawson, P.: Individuals: an essay in descriptive metaphysics. Routledge (1959)
[64] Damasio, A.: Looking for Spinoza. Putnam (2003)