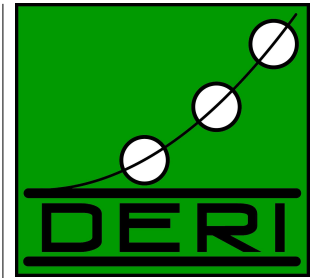


DERI – DIGITAL ENTERPRISE RESEARCH INSTITUTE



OWL DL vs. OWL FLIGHT:  
CONCEPTUAL MODELING AND  
REASONING FOR THE SEMANTIC  
WEB

Jos De Bruijn    Axel Polleres    Rubén Lara  
                         Dieter Fensel

DERI TECHNICAL REPORT 2004-11-10  
NOVEMBER 2004

DERI – DIGITAL ENTERPRISE RESEARCH INSTITUTE

**DERI Ireland**  
University Road  
Galway  
IRELAND  
[www.deri.ie](http://www.deri.ie)

**DERI Innsbruck**  
Technikerstrasse 13  
A-6020 Innsbruck  
AUSTRIA  
[www.deri.ie](http://www.deri.ie)



## OWL DL vs. OWL FLIGHT: CONCEPTUAL MODELING AND REASONING FOR THE SEMANTIC WEB

Jos De Bruijn<sup>1</sup>, Axel Polleres<sup>1</sup>, Rubén Lara<sup>1</sup>, Dieter Fensel<sup>1</sup>

**Abstract.** The Semantic Web languages RDFS and OWL have been around for some time now. However, the presence of these languages has not brought the breakthrough of the Semantic Web the creators of the languages had hoped for. OWL has a number of problems in the area of interoperability and usability in the context of many practical application scenarios which impede the connection to the Software Engineering and Database communities. In this paper we present OWL Flight, which is loosely based on OWL, but the semantics is grounded in Logic Programming rather than Description Logics, and it borrows the constraint-based modeling style common in databases. This results in different types of modeling primitives and enforces a different style of ontology modeling. In this paper we analyze the modeling paradigms of OWL DL and OWL Flight, as well as reasoning tasks supported by both languages. We argue that different applications on the Semantic Web require different styles of modeling and thus both types of languages are required for the Semantic Web.

**Keywords:** Semantic Web, Ontologies, Description Logics, Logic Programming.

---

<sup>1</sup>Digital Enterprise Research Institute Innsbruck, E-mail: {jos.debruijn, axel.polleres, ruben.lara, dieter.fensel}@deri.org.

**Acknowledgements:** The work is funded by the European Commission under the projects DIP, Knowledge Web, InfraWebs, SEKT, SWWS, ASG and Esperonto; by Science Foundation Ireland under the DERI-Lion project; by the Vienna city government under the CoOperate programme and by the FIT-IT (Forschung, Innovation, Technologie - Informationstechnologie) under the projects RW<sup>2</sup> and TSC.

Copyright © 2004 by the authors

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Restrictions and constraints</b>	<b>2</b>
<b>3</b>	<b>OWL DL</b>	<b>4</b>
<b>4</b>	<b>Pitfalls of OWL</b>	<b>6</b>
4.1	Interoperability . . . . .	6
4.2	Scalability . . . . .	7
4.3	Conceptual Modeling . . . . .	7
4.4	Extensibility . . . . .	9
<b>5</b>	<b>OWL Flight</b>	<b>10</b>
5.1	OWL Flight Abstract Syntax and Mapping to F-Logic . . . . .	11
<b>6</b>	<b>Conceptual Modeling on the Semantic Web</b>	<b>13</b>
6.1	Cardinality Restrictions and Constraints . . . . .	14
6.2	Value restrictions vs. constraints . . . . .	14
6.3	Meta-modeling . . . . .	15
6.4	Complete Class Definitions . . . . .	16
6.5	Negation . . . . .	17
<b>7</b>	<b>Reasoning tasks on the Semantic Web</b>	<b>17</b>
7.1	Subsumption . . . . .	18
7.2	Query Answering . . . . .	19
<b>8</b>	<b>Conclusions</b>	<b>20</b>

## 1 Introduction

The vision of the Semantic Web [Berners-Lee et al., 2001] is to enable automatic interoperation between entities on the Web. Such interoperation can be achieved through annotation of the content on the Web with machine-processable data. When such annotations are linked to ontologies, machines can achieve a certain degree of understanding of the data. Ontologies [Fensel, 2003] are formal and explicit specifications of certain domains and are shared between large groups of stakeholders. These properties make ontologies ideal for machine processing and enabling interoperation. In fact, ontologies form the backbone of the Semantic Web and are the key to enable automated interoperation and collaboration. An ontology typically consists of a number of classes, a number of relations (sometimes called properties) between these classes, a number of instances and a number of axioms. These elements are all expressed using some logical language.

In order to allow sharing and reuse of ontologies on the Semantic Web, a common ontology language is required. The W3C has developed two ontology languages for use on the Semantic Web. The first is RDFS [Brickley and Guha, 2004], which was developed as a lightweight ontology language. The second language is OWL [Dean and Schreiber, 2004], which is a more expressive ontology language based on Description Logics [Baader et al., 2003].

OWL consists of three species, namely OWL Lite, OWL DL and OWL Full, which are intended to be layered according to increasing expressiveness. OWL Lite is a notational variant of the Description Logic  $\mathcal{SHIF}(\mathbf{D})$ ; OWL DL is a notational variant of the Description logic  $\mathcal{SHOIN}(\mathbf{D})$  [Horrocks and Patel-Schneider, 2003]. It turns out that OWL DL adds very little in expressiveness to OWL Lite [Horrocks et al., 2003]. OWL Lite and OWL DL pose several restrictions on the use of RDF and redefine the semantics of the RDFS primitives; thus, OWL Lite and OWL DL are not properly layered on top of RDFS. The most expressive species of OWL, OWL Full, layers on top of both RDFS and OWL DL, and because these languages are so different, the semantics of OWL Full are not straightforward and are not a proper extension of the OWL DL semantics (see also Section 4.1). The lack of proper layering between RDFS and the less expressive species of OWL and the lack of proper layering between OWL DL and OWL Lite on the one side and OWL Full on the other, raises doubts about interoperability between ontologies written in these different languages.

Besides the lack of interoperability between the OWL species, there are some pitfalls in ontology modeling with OWL for certain domains. OWL adheres to the *open-world assumption* (OWA), which means that if a certain statement is not known to be true, the negation of the statement can not be inferred. In contrast, databases adhere to the *closed world assumption* (CWA), which means that if a certain statement is not known to be true, it is assumed to be false, i.e. its negation is true. Thus, databases implicitly assume complete knowledge. Furthermore, OWL does not adhere to the *unique name assumption* (UNA; common in databases and software engineering), which means that from the fact that two individuals have different names, it does not follow that these names denote different individuals. In contrast, databases do adhere to the unique name assumption, which means that each identifier identifies a distinct individual.

The closed world assumption and the unique name assumption are common in Software Engineering and Database Systems. For computing professionals from the areas of Software Engineering and Database Systems, there are some modeling pitfalls in OWL (see also [de Bruijn et al., 2004b]):

- Behavior of cardinality restrictions (the allowed number of values for a property): equality between individuals or existence of individuals outside the knowledge base can be inferred. The cardinality of properties is not *checked*, but *inferred*.

- Behavior of range restrictions: the type of property values can be inferred. The type of property values is not *checked*, but *inferred*.

Some of these potential pitfalls can be eliminated by introducing the unique name assumption and asserting complete knowledge about certain descriptions. The unique name assumption can be introduced in a Description Logic knowledge base by asserting inequality between each distinct pair of individual names. OWL DL offers the `allDifferent` operator for this purpose. For modeling complete knowledge, the Description Logic community offers the epistemic operator **K** [Donini et al., 1998]. When using the **K** operator as a prefix to a description, the description represents complete knowledge about the description, i.e. the knowledge base is assumed to contain all instances of this description.

In this paper we are not concerned with possible extensions of OWL or with the use of explicit assertions in order to enforce certain behavior of the modeling primitives. We are concerned with OWL DL as it is and with the modeling constructs provided by the language. In order to assert the unique name assumption in an OWL DL knowledge base, it is necessary to include all individuals in the knowledge base in the `allDifferent` statement. This is not feasible in practice. Clearly, OWL could be extended with a construct with the implicit meaning that all individuals are different and OWL could also be extended with a construct which asserts complete knowledge for a certain description. However, we examine in this paper the OWL language as it is and the properties of the OWL modeling constructs as they are being used today.

In order to investigate these modeling pitfalls of OWL and the usability of OWL for modeling and reasoning on the Semantic Web, we describe in this paper OWL Flight [de Bruijn et al., 2004c], an ontology language based on the Logic Programming subset of OWL [de Bruijn et al., 2004b] which is inspired by the intersection of Logic Programming and Description Logic [Grosz et al., 2003] with certain extensions in the area of datatypes [Pan and Horrocks, 2004], database-style constraints and meta-modeling. Our two main motivations for creating OWL Flight were: (1) eliminate some of the pitfalls in conceptual modeling with OWL and (2) enable efficient query answering using common off-the-shelf reasoning engines which benefit from many years of research on optimizing query answering for deductive databases (e.g. [Ramakrishnan, 1991]).

This paper is further structured as follows. We first discuss the formal and conceptual differences between restrictions and constraints in Section 2. Then, we briefly describe OWL DL and pitfalls in the use of this language for the Semantic Web in Sections 3 and 4, followed by the introduction of OWL Flight in Section 5. We analyze the conceptual modeling features of OWL DL and OWL Flight in the light of potential modeling tasks on the Semantic Web in Section 6. We contrast the reasoning tasks supported by the languages and the use cases for these reasoning tasks in Section 7. Finally, we provide some conclusions and mention future work in Section 8.

## 2 Restrictions and constraints

In order to provide a better understanding of modeling and reasoning with different types of languages, we describe the difference between restrictions and constraints. The terms *restriction* and *constraint* both refer to some aspects of a property, such as the cardinality or the range of the property. We describe the differences first from a formal logical point of view and then from a conceptual modeling point of view.

For the purposes of our formal treatment of restrictions and constraints, we see an ontology as a logical theory where class definitions, property definitions, etc. are formulae in the theory. From a logical point of view, a restriction applied to a class definition can be seen as a first-order formula. Thus, a restriction *restricts* the number of models of a logical theory. Restrictions are thus an integral part of the logical theory. For an ontology  $\Omega$  and a restriction  $r$ , let  $M^\Omega$  be the set of models of  $\Omega$  and  $M^{\Omega \cup r}$  be the set of models of  $\Omega \cup r$ , then  $M^\Omega \supseteq M^{\Omega \cup r}$ . Let  $cons(\Omega)$  be the set of consequences of  $\Omega$  and  $cons(\Omega \cup r)$  be the set of consequences of  $\Omega \cup r$ , then  $cons(\Omega) \subseteq cons(\Omega \cup r)$ . Thus, restrictions allow to *infer* additional information, because they increase the number of consequences.

Constraints<sup>1</sup> specify conditions which may not be violated by an interpretation of the logical theory. Constraints are not part of the logical theory and thus they do not increase the number of consequences of the theory. For ontology  $\Omega$  and constraint  $c$ , if  $\Omega \not\models c$  we say that the constraint is violated. However, since we don't see  $c$  as part of the theory, it does not affect the set of consequences. In summary, constraints do not allow to infer additional information, instead, they can be used to *check* the knowledge base with respect to certain conditions.

From a modeling point of view, restrictions and constraints are both used to specify certain aspects of a property, namely cardinality and range.

The *cardinality* of a property is the number of values a particular property may have for a particular individual. It is possible to specify a minimal bound and a maximal bound on the cardinality which we will refer to as minimal and maximal cardinality, respectively. Both minimal and maximal cardinality can be an arbitrary positive integer. Say we have a property  $P$  and an individual  $a$ . In case  $P$  has a minimal cardinality restriction of  $n$ , and the number of values of  $P$  for  $a$  in the knowledge base is smaller than  $n$ , a number of unknown individuals is inferred to exist outside of the knowledge base. In case  $P$  has a maximal cardinality restriction of  $n$ , and there are more than  $n$  values of  $P$  for  $a$  in the knowledge base, equality between property values is inferred in order to make the knowledge fit the restriction.

As for a minimal (or maximal, respectively) cardinality *constraint* of  $n$  on  $P$ , whenever the number of different values of  $P$  for  $a$  derivable from the knowledge base is smaller (or greater, respectively) than  $n$ , the constraint is violated and thus the knowledge base is erroneous.

Hence, cardinality restrictions can be used to *infer* existence of unknown individuals or equality between known individuals and constraints can be used to *check* available knowledge with respect to the constraints and do not allow for additional inference.

The *range* of a property is the type a property value may have. Say we have a property  $P$  with a range restriction  $C$  and a tuple  $\langle a, b \rangle \in P$ . In case it is not known that  $b$  is of type  $C$ , it is inferred that  $b$  is of type  $C$ . In contrast, if  $P$  would have a range constraint  $C$ , then from the fact that it is not known that  $b$  is of type  $C$ , the constraint is violated and the knowledge base is inconsistent. Hence, range restrictions can be used to *infer* (presumably missing) type information about individuals and range constraints can be used to check available type information with respect to the constraint.

From a conceptual modeling point of view, the term "restriction" might be misleading, because it might suggest that the number of consequences of the theory is restricted, whereas actually the number of models is restricted and thus the number of consequences is increased. In the extreme case where the number of models is restricted to 0 and thus the theory is inconsistent, each sentence

---

<sup>1</sup>Our notion of constraints is similar to the notion of integrity constraints in logic programming and databases.

is actually a consequence of the theory, by the explosive nature of classical logic.

### 3 OWL DL

In this paper we are mainly concerned with the most well-known and most investigated species of OWL, namely OWL DL, which can be seen as an alternate notation for the Description Logic language  $\mathcal{SHOIN}(\mathbf{D})$  [Horrocks and Patel-Schneider, 2003].

OWL DL has different syntaxes, the most prominent being the RDF/XML syntax, which is actually used in the language reference. However, the normative syntax for OWL DL is the abstract syntax, described in [Patel-Schneider et al., 2004], which we will use for the examples in the remainder of this paper for reasons of legibility. In the remainder of this section we will explain OWL DL using Description Logic syntax. See Tables 1 and 2 for a mapping between the OWL DL abstract syntax and the syntax of the Description Logic  $\mathcal{SHOIN}(\mathbf{D})$ .

A Description Logic knowledge base consists of two parts, namely the TBox and the ABox. The TBox consists of a number of class and property axioms; the ABox consists of a number of individual assertions (see Table 1). Here, letters  $C, D$  refer to class names,  $T$  refers to a concrete datatype, whereas  $R$  refers to an object property name,  $U$  refers to datatype property;  $Q$  refers to an object or datatype property where several appearances of  $Q_i, Q_j$  in one statement always refer to either both object or both datatype properties;  $o$  and  $t$  refer to object and concrete values, respectively. A class axiom in the TBox consists of two class descriptions, separated with the GCI (General Class Inclusion, or subsumption;  $\sqsubseteq$ ) symbol or the equivalence symbol ( $\equiv$ ), which is equivalent to GCI in both direction (i.e.  $\sqsubseteq$  and  $\sqsupseteq$ ). Similarly, a property axiom consists of a two property names, separated with the subsumption ( $\sqsubseteq$ ) or the equivalence ( $\equiv$ ) symbol. A description in the TBox is either a named class ( $A$ ), an enumeration ( $\{o_1, \dots, o_n\}$ ), an property restriction ( $\exists R.D, \forall R.D, \exists R.o, \geq nR, \leq nR$ , analogously for datatype property restrictions), or an intersection ( $C \sqcap D$ ), union ( $C \sqcup D$ ) or complement ( $\neg C$ ) of such descriptions (Table 2). Individual assertions in the ABox are either class membership ( $o \in C_i$ ), property value ( $\langle o_1, o_2 \rangle \in R_i, \langle o_1, t_1 \rangle \in U_i$ ), or individual (in)equality ( $o_1 = o_2, o_1 \neq o_2$ ) assertions (Table 1).

Description Logics make the distinction between abstract and concrete properties, based on whether the range of the property is in the abstract or the concrete domain. OWL DL reflects this distinction by distinguishing object properties and datatype properties, where an object property may only have a description as its range and a datatype property may only have a datatype as its range; descriptions and datatypes are disjoint.

OWL DL has a direct model-theoretic semantics [Patel-Schneider et al., 2004] similar to the model-theoretic semantics for Description Logics [Baader et al., 2003]. It was shown that entailment in OWL DL (checking whether one ontology is logically entailed by another) can be reduced to satisfiability checking in Description Logics [Horrocks and Patel-Schneider, 2003]. Since all interesting reasoning tasks can be reduced to Description Logic (un)satisfiability, it is safe to say that OWL DL is a notational variant of a Description Logic. Furthermore, it was shown that Description Logics are a subset of First-Order Logic [Borgida, 1996], thus OWL DL is also a notational variant for a subset of First-Order Logic.

Description Logics have a set-based model-theoretic semantics. In an interpretation  $\mathcal{I}$ , a description  $C$  (Table 2) is mapped to a subset of the domain  $\Delta^{\mathcal{I}}$  and an individual  $o$  is mapped to an object of  $\Delta^{\mathcal{I}}$  using the mapping function  $\cdot^{\mathcal{I}}$ . Similarly, a datatype  $T$  is mapped to a subset of the concrete domain  $\Delta_D^{\mathcal{I}}$  and a literal is mapped to a value in  $\Delta_D^{\mathcal{I}}$ . An abstract role  $R$  is mapped

OWL Abstract Syntax	DL syntax
<i>Class axioms</i>	
Class( <i>A</i> partial $C_1 \dots C_n$ )	$A \sqsubseteq C_i$
Class( <i>A</i> complete $C_1 \dots C_n$ )	$A \equiv C_1 \sqcap \dots \sqcap C_n$
EnumeratedClass( <i>A</i> $o_1 \dots o_n$ )	$A \equiv \{o_1, \dots, o_n\}$
SubClassOf( $C_1$ $C_2$ )	$C_1 \sqsubseteq C_2$
EquivalentClasses( $C_1 \dots C_n$ )	$C_1 \equiv \dots \equiv C_n$
DisjointClasses( $C_1 \dots C_n$ )	$C_i \sqcap C_j \sqsubseteq \perp$
<i>Property axioms</i>	
ObjectProperty( <i>R</i> )	
super( $R_1 \dots R_n$ )	$R \sqsubseteq R_i$
domain( $C_1 \dots C_n$ )	$\top \sqsubseteq \forall R^-.C_i$
range( $C_1 \dots C_n$ )	$\top \sqsubseteq \forall R.C_i$
[inverseOf( $R_0$ )]	$R \equiv R_0^-$
[Symmetric]	$R \equiv R^-$
[Functional]	$\top \sqsubseteq \leq 1R$
[InverseFunctional]	$\top \sqsubseteq \leq 1R^-$
[Transitive])	Trans( $R$ )
Datatype( <i>T</i> )	
DatatypeProperty( <i>U</i> )	
super( $U_1 \dots U_n$ )	$U \sqsubseteq U_i$
domain( $C_1 \dots C_n$ )	$\top \sqsubseteq \forall U^-.C_i$
range( $T_1 \dots T_n$ )	$\top \sqsubseteq \forall U.T_i$
[Functional])	$\top \sqsubseteq \leq 1U$
SubPropertyOf( $Q_1$ $Q_2$ )	$Q_1 \sqsubseteq Q_2$
EquivalentProperties( $Q_1 \dots Q_n$ )	$Q_1 \equiv \dots \equiv Q_n$
<i>Individual assertions</i>	
Individual( <i>o</i> )	
type( $C_1 \dots C_n$ )	$o \in C_i$
value( $R_1$ $o_1$ ) ... value( $R_m$ $o_m$ )	$\langle o, o_i \rangle \in Q_i$
value( $U_1$ $t_1$ ) ... value( $U_n$ $t_n$ )	$\langle o, t_i \rangle \in U_i$
SameIndividual( $o_1 \dots o_n$ )	$o_1 = \dots = o_n$
DifferentIndividuals( $o_1 \dots o_n$ )	$o_i \neq o_j, i \neq j$

Table 1: Axioms in OWL DL and  $\mathcal{SHOIN}(\mathbf{D})$ 

to a binary relation over the abstract domain domain:  $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ . Similarly, a concrete role  $R$  is mapped to a binary relation between the abstract and the concrete domain:  $\Delta^{\mathcal{I}} \times \Delta_D^{\mathcal{I}}$ . Equivalence of descriptions is interpreted as set equivalence ( $C \equiv D$  is interpreted as  $C^{\mathcal{I}} = D^{\mathcal{I}}$ ), subsumption is interpreted as set inclusion ( $C \sqsubseteq D$  is interpreted as  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ ), and so on. We refer the interested reader to [Baader et al., 2003, Chapter 2] for a more exhaustive treatment of Description Logic semantics.

There exist several implementations for reasoning with Description Logics (e.g. FaCT++ [Tsarkov and Horrocks, 2002], RACER [Haarslev and Möller, 2003]) which implement different reasoning tasks in Description Logic languages. Two important reasoning tasks in Description Logics are subsumption checking and checking class membership [Baader et al., 2003]. Subsumption checking amounts to checking whether one class is a subclass of another concept, i.e. checking whether one concept is more specific than another concept. The class membership inference is used to check whether an individual is a

OWL Abstract Syntax	DL syntax
$A$ (URI Reference)	$A$
<code>owl:Thing</code>	$\top$
<code>owl:Nothing</code>	$\perp$
<code>intersectionOf(<math>C_1 \dots C_n</math>)</code>	$C_1 \sqcap \dots \sqcap C_n$
<code>unionOf(<math>C_1 \dots C_n</math>)</code>	$C_1 \sqcup \dots \sqcup C_n$
<code>complementOf(<math>C</math>)</code>	$\neg C$
<code>oneOf(<math>o_1 \dots o_n</math>)</code>	$\{o_1, \dots, o_n\}$
<code>restriction(<math>R</math> someValuesFrom(<math>C</math>))</code>	$\exists R.D$
<code>restriction(<math>R</math> allValuesFrom(<math>C</math>))</code>	$\forall R.D$
<code>restriction(<math>R</math> value(<math>o</math>))</code>	$\exists R.o$
<code>restriction(<math>R</math> minCardinality(<math>n</math>))</code>	$\geq nR$
<code>restriction(<math>R</math> maxCardinality(<math>n</math>))</code>	$\leq nR$
<code>restriction(<math>U</math> someValuesFrom(<math>T</math>))</code>	$\exists U.T$
<code>restriction(<math>U</math> allValuesFrom(<math>T</math>))</code>	$\forall U.T$
<code>restriction(<math>U</math> value(<math>t</math>))</code>	$\exists U.t$
<code>restriction(<math>U</math> minCardinality(<math>n</math>))</code>	$\geq nU$
<code>restriction(<math>U</math> maxCardinality(<math>n</math>))</code>	$\leq nU$

Table 2: Descriptions in OWL DL and *SHOIN*

member of a specific class.

## 4 Pitfalls of OWL

In this section we describe a number of (potential) pitfalls of OWL with respect to interoperability on the Semantic Web and scalability of reasoning with the language. Furthermore, we discuss the suitability of its modeling constructs and modeling style for certain domains and extensibility of the language in the direction of rules. The rule language was identified as the next layer in the Semantic Web language stack and extends the ontology language with the possibility to write arbitrary rules.

### 4.1 Interoperability

Problems of interoperability might occur between RDFS and OWL and between the species of OWL because of problems in the layering of the languages.

The OWL language is layered on top of RDFS. However, only the most expressive species of OWL, namely OWL Full, is completely syntactically and semantically layered on top of RDFS. The less expressive species of OWL, namely OWL Lite and OWL DL, pose syntactical restrictions on the use of RDF and redefine the semantics of the RDFS modeling primitives.

The species of OWL are layered according to increasing expressiveness, where OWL Lite is the least expressive and OWL Full the most expressive. On the one hand, OWL Lite poses many syntactical restrictions on the constructs which can be used in ontology modeling. On the other hand, the only feature really added by OWL DL compared with OWL Lite is the use of *nominals* (individuals in class descriptions) [Horrocks et al., 2003]; all other features of OWL DL can be written down using OWL Lite through complicated syntactical constructions. The third, and most expressive, species of OWL, namely OWL Full, is unfortunately not properly semantically

layered on top of OWL DL; entailment under OWL DL semantics is not equivalent to entailment under OWL Full semantics for the same ontology: OWL Full allows additional inferences. This discrepancy is caused by the incompatibility between the model-theoretic semantics of OWL DL and the axiomatic semantics of and syntactical freedom of RDFS. This raises doubts about the level of interoperability between the different species of OWL.

Both the layering of OWL on top of RDFS and the layering of the OWL species (especially the layering of OWL Full on top of OWL DL) is, in our opinion, inappropriate. The two less expressive species of OWL, OWL Lite and OWL DL, are only layered on top of a restricted subset of RDFS, whereas the most expressive species of OWL, OWL Full, is completely syntactically and semantically layering on top of RDFS, but not on top of OWL Lite and OWL DL. This improper layering might hamper interoperability between agents using RDFS or OWL Full on the one side and agents using OWL Lite or OWL DL on the other. In fact, because the main aim of the Semantic Web is to enable interoperability on a world-wide scale, the lack of interoperability between Semantic Web languages can seriously hamper the development of the Semantic Web.

## 4.2 Scalability

There are doubts with respect to the scalability of certain reasoning tasks in OWL, mostly query answering. Description Logic reasoning and optimization has so far mostly focused on the optimization of the subsumption inference; few optimizations exist for query answering. Optimizing query answering for Description Logics is currently still very much a research issue [Grosf et al., 2003, Hustadt et al., 2004, Haarslev and Möller, 2004, Horrocks et al., 2004a].

The satisfiability problem in *SHOIN*, the Description Logic underlying OWL DL, has NExp-Time (non-deterministic exponential time) worst-case complexity. Most current Description Logic implementations use complex Tableaux satisfiability checks for query answering [Horrocks et al., 2000]. The major problem with using Tableaux for query answering for knowledge bases with more than a few individuals is that a Tableaux check is required for each individual in the knowledge base to check whether it is in the answer to the query. Although there exist several optimizations [Haarslev and Möller, 2004, Horrocks et al., 2004a], the fundamental problems are not solved. In our opinion it is a mistake to require such complex reasoning for even the least expressive of the OWL species, since efficient ABox reasoning will most likely play a major role on the Semantic Web. There are currently over four billion web pages indexed by Google, thus, in order for the Semantic Web to work outside of the research lab, it must be possible to reason with large collections of instances.

On the contrary, there is a large body of research work on query optimization in the context of deductive databases (e.g. [Ramakrishnan, 1991]). The Semantic Web could benefit from this research by adopting reasoning mechanisms from the deductive database area and reusing existing reasoner implementations. However, the Deductive Databases and Description Logics paradigms are incompatible and deductive databases can only be used for query answering with a restriction subset of Description Logics (cf. [Grosf et al., 2003]).

## 4.3 Conceptual Modeling

Some of the modeling constructs of OWL DL have a semantics which might seem odd to people not familiar with Description Logics. These constructs concern the treatment of abstract vs. concrete

values, cardinality restrictions, and value restrictions. Furthermore, we identify limitations in the support for datatypes in OWL.

**Difference in the treatment of abstract and concrete values** The semantics of OWL is defined through an interpretation  $\mathcal{I}$  over two distinct domains, namely the abstract domain  $\Delta^{\mathcal{I}}$ , which consists of all individuals, and the datatype domain  $\Delta_D^{\mathcal{I}}$ , which consists of all data values.  $\Delta^{\mathcal{I}}$  and  $\Delta_D^{\mathcal{I}}$  are disjoint. This disjointness requires a different treatment of values from on the one hand the abstract domain and on the other hand the concrete domain. Reasoning over the concrete domain is done through an external datatype oracle.

Literals (concrete values) in OWL, other than abstract individuals, adhere to the *unique name assumption*. Furthermore, the closed-world assumption is applied to the concrete domain. Consequently, cardinality and range restrictions exhibit different behavior depending on whether they are concerned with object or datatype properties. Restrictions over object properties exhibit the usual behavior of restriction as we have described in Section 2. Restrictions over datatype properties, however, exhibit the behavior of constraints as described in Section 2. It has been argued that the difference in treatment of individuals and literals makes sense, because the domain of a data type is known. We argue that from the point of view of the user of the language, this conceptual distinction is not intuitive. Whereas restrictions involving the abstract domain are used to *infer* new knowledge, restrictions involving the concrete domain are used to *check* whether the knowledge satisfies certain constraints.

**Deriving Equality through Cardinality Restrictions** In OWL, it is possible to specify a maximal cardinality restriction for a property. When there are more instances of this property having the same domain value, than the maximal cardinality restriction prescribes, equality between individuals is inferred. We illustrate this with an example.

**Example 4.1** Assume the following OWL DL knowledge base:

```
ObjectProperty(hasPassenger domain(FlightSeat)
  range(Passenger))
Class(FlightSeat partial
  restriction(hasPassenger maxCardinality(1)))
Individual(seat1 type(FlightSeat)
  value(hasPassenger mary)
  value(hasPassenger john))
```

`FlightSeat` represents the seats in a particular flight. The property `hasPassenger` associates a seat in a flight with a passenger. A seat may only have one passenger, which is guaranteed by the restriction `maxCardinality(1)`. The individual `seat1` is asserted as instance of `FlightSeat`; `seat1` has two values for the property `hasPassenger`, namely `mary` and `john`.

From the OWL knowledge base in Example 4.1 the reasoner will draw the conclusion that `mary` and `john` both refer to the same passenger. The possibility that there was a mistake in the system (either at the ontology or at the instance level) is not taken into account. This could lead to the booking of several passengers on a single seat, unless the unique name assumption is enforced.

**Deriving Class Membership through Value Restrictions** In OWL it is possible to restrict the range of a property  $P$  to a class description  $C$  either through a range restriction in the property definition or through a local universal range restriction in a class definition. From this restriction it is inferred that every value of this property is a member of class  $C$ .

**Example 4.2** We take the OWL knowledge base from Example 4.1 and add the following assertions:

```
Individual(seat3 type(FlightSeat))
Individual(seat2 type(FlightSeat)
  value(hasPassenger seat3))
```

The above assertions introduces two new instances of the class *FlightSeat* by the names of *seat2* and *seat3* plus a value for the property *hasPassenger* at *seat3*, namely *seat3*.

From Example 4.2 we can infer that *seat3* is a *Passenger*. Clearly, there is some mistake in the individual assertions, because a seat cannot occupy another seat; only a passenger can. However, this mistake is not detected; instead the modeling mistake allows for additional (incorrect) inferences. Although it is possible in OWL DL to express disjointness of classes, in which case an inconsistency would be derived, we argue that in many application domains it is natural to assume disjointness of classes beforehand and only deviate from this assumption when classes are *known* not to be disjoint.

**Meta-modeling** Meta-modeling, and especially the treatment of classes as instances, has been identified as an important style of modeling for the Semantic Web (cf. [Schreiber, 2002, Noy et al., 2004]). RDFS and OWL Full support meta-modeling. However, the entailment problem in OWL Full is undecidable in general and RDFS is not expressive enough for many purposes. Meta-modeling is not supported in OWL DL and is thus not available for most practical Semantic Web applications.

**Limited Support for Datatypes** OWL allows for a limited treatment of datatypes. Only unary datatypes are supported by OWL. The three major limitations of datatype support in OWL are [Pan and Horrocks, 2004]: (1) lack of negated datatypes, which is required for most Description Logic reasoners, (2) lack of support for datatype predicates; it is only possible to refer to a single value in a datatype domain and not, e.g. to express the greater-than ( $\geq$ ) relation for the `xsd:integer` domain, and (3) lack of support for user-defined datatypes.

OWL-E [Pan and Horrocks, 2004] is an extension of OWL with so-called datatype groups. Datatype groups overcome the aforementioned limitations of datatype support in OWL and bridge the gap between datatypes in OWL and concrete domains as they have been investigated in the Description Logic community (see e.g. [Horrocks and Sattler, 2001]).

#### 4.4 Extensibility

It has been suggested that besides an ontology language, the Semantic Web needs a rule language. Such a rule language would provide additional expressiveness on top of the ontology language. It was identified as a requirement that the rule language is an extension of the ontology language.

There have been several proposals for rule extensions of Description Logic languages (e.g. CARIN [Levy and Rousset, 1998] and SWRL [Horrocks et al., 2004b]). Three general directions

can be identified in these approaches: (1) approaches which allow the use of classes (and properties) as unary (binary) predicates in the body of the rules (e.g. CARIN [Levy and Rousset, 1998]) (2) approaches which directly extend the Description Logic knowledge base with Horn-style rules (e.g. SWRL [Horrocks et al., 2004b]) and (3) approaches which define a restricted interface between the Description Logic and the Logic Programming formalisms and allow for exchange of conclusions between the two (e.g. [Eiter et al., 2004]). The approaches under (1) and (2) are similar, however, the approaches under (1) do not allow predicates from the Description Logic knowledge base in the heads of the rules, whereas the approaches under (2) do, which means that in these approaches conclusions drawn from the rules can affect the Description Logic knowledge base.

SWRL (Semantic Web Rule Language) has been proposed as a rule language for the Semantic Web, layered on top of OWL DL<sup>2</sup>. Satisfiability of an OWL DL knowledge base augmented with SWRL rules is undecidable, as was pointed out by the authors of the proposal [Horrocks et al., 2004b]. Straightforward extension of Description Logics with rules leads to undecidability [Levy and Rousset, 1998] and does not allow the use of existing rule systems to reason with the language. Instead, complex new calculi need to be developed, or first-order theorem proving is required.

## 5 OWL Flight

As opposed to OWL DL, we will in this section define a novel variant of OWL, called OWL Flight, which addresses some of the above-mentioned problems. On the one hand, OWL Flight restricts the OWL syntax such that it falls in the Datalog fragment and thus query answering can be done using a Logic Programming implementation (cf. [Grosz et al., 2003]). On the other hand, we take this restricted subset of OWL DL as a basis which we further extend in order to overcome some of the limitations of OWL DL. These extensions include integrity constraints, a more elaborate treatment of datatypes than OWL based on OWL-E [Pan and Horrocks, 2004], and meta-modeling features where we make use of the meta-modeling features of the Datalog compatible variant of F-Logic [Kifer et al., 1995].

The formalism underlying OWL Flight is Datalog<sup>IC,≠,not</sup>, which is Datalog extended with integrity constraints, inequality and default negation. OWL Flight is an extension of the so called DLP fragment of OWL which marks the intersection of Description Logics and Datalog [Grosz et al., 2003, de Bruijn et al., 2004b]. Compared to DLP, we further add limited support for nominals, as well as a meta-modeling facility (e.g. treating classes as instances) and constraints.

Features of OWL DL not included in OWL Flight are: enumerated classes, individual (in)equality assertions, complements, property restrictions in complete class definitions. Furthermore, we restrict the use of property restrictions, nominals, and union on the left- and right-hand-side of the GCI.

OWL Flight adds the following distinct features to this restricted subset of OWL DL: We adopt the *unique name assumption*<sup>3</sup> and adding *cardinality constraints*, i.e. checking cardinali-

<sup>2</sup>We are not aware of any attempts to create a rule language based on OWL Full. However, such a language would certainly be undecidable, as OWL Full itself is already undecidable.

<sup>3</sup>Notice that also the prominent Description Logic reasoner RACER assumes unique names, not implementing equality reasoning, for efficiency reasons. As pointed out above, we do not necessarily see the lack of equality reasoning as a drawback but even as an advantage, due to unintuitive effects it might have with respect to reasoning in OWL DL.

ties rather than the non-intuitive inference of equality. *Property value constraints* eliminate the non-intuitive inferring of class membership. Constraints adopt the closed world assumption, allowing to check the data in the knowledge base, i.e. a single ontology (along with imported ontologies); as mentioned above, this feature is useful in many contexts and missing in OWL DL (cf. [Heflin and Muñoz-Avila, 2002, Donini et al., 1998]). *Meta-modeling* returns modeling support lost in the transition from RDFS to OWL DL. *More elaborate treatment of datatypes*, following OWL-E [Pan and Horrocks, 2004], which overcomes the limitations of the treatment of datatypes in OWL.

### 5.1 OWL Flight Abstract Syntax and Mapping to F-Logic

Syntactically, we base OWL Flight on a subset of the abstract syntax of OWL DL, but add distinct constructs for constraints and eliminate the separation of the vocabulary. Table 3 shows a feature comparison of OWL DL and OWL Flight based on an extended version of the OWL abstract syntax. The meaning of the letters in the table corresponds with the description given in Section 3; furthermore,  $E$  stands for a datatype predicate and  $n$  stands for the arity of the datatype predicate. The terms *lhs* (and *rhs*, resp.) in the table express that certain descriptions in OWL Flight are restricted to be used in the left-hand side (or right-hand side, resp.) side. Descriptions allowed on the right-hand side are allowed in **partial** class definitions and in the second argument of **SubClassOf**. Descriptions allowed on the left-hand side are allowed in the first argument of **SubClassOf**. Descriptions allowed on both sides are also allowed in **complete** class definitions. On the one hand, unrestricted usage of these features would lead us outside the expressivity of Datalog<sup>IC,≠,not</sup>. On the other hand, some of the unintuitive features of OWL DL like inferring equality originate precisely from this unrestricted usage. In the lower part of Table 3 we find the features newly added in OWL Flight, which are constraints in partial class definitions and Datatype expressions. As described in Section 2, constraints check all inferred instances of the respective class on the *lhs* for the condition on the *rhs*. Datatype expressions follow the idea of [Pan and Horrocks, 2004], which allows to define  $n$ -ary datatype predicates which constrain the values of the tuples given by datatype properties  $U_1, \dots, U_n$  of an object. So, with this extension, one can express relations between datatype properties, for instance that the number of bookings for a hotel room has to be smaller than its capacity, etc.

Note that constraints are only allowed for object properties, because for datatype properties the **restriction** keyword in OWL DL already has a constraining semantics as discussed in Section 4. For the sake of brevity, we refer to [de Bruijn et al., 2004c] for a complete description of the abstract syntax of OWL Flight.

We define the semantics of OWL Flight through a mapping to the Datalog subset of F-Logic [Kifer et al., 1995]. By doing so, we gain the following benefits: Staying within the Datalog world, we can directly use implemented engines tailored for efficient query answering. F-Logic syntax with its origins in Frame based modeling (as opposed to pure Datalog) directly allows for meta-modeling features and provides constructs for class membership, subclassing and attributes in the language, rather than relying on predicates to axiomatize the behavior of these constructs. The complete mapping is defined in [de Bruijn et al., 2004c] and we restrict ourselves to the rough idea in this paper.

Let  $x, y, z$  be variables,  $:$  stands for class membership,  $::$  stands for the subclass relationship and a molecule of the form  $A[B \rightarrow C]$  stands for “object  $A$  has an attribute  $B$  with value  $C$ ”. Such

Abstract Syntax	DL	Flight
<i>Axioms and Individual Assertions</i>		
Class( <i>A</i> partial $C_1 \dots C_n$ )	+	+
Class( <i>A</i> complete $C_1 \dots C_n$ )	+	+
EnumeratedClass( <i>A</i> $o_1 \dots o_n$ )	+	-
SubClassOf( $C_1 C_2$ )	+	+
EquivalentClasses( $C_1 \dots C_n$ )	+	+
DisjointClasses( $C_1 \dots C_n$ )	+	+
ObjectProperty( <i>R</i> ...)	+	+
Datatype( <i>T</i> )	+	+
DatatypeProperty( <i>U</i> ...)	+	+
SubPropertyOf( $Q_1 Q_2$ )	+	+
EquivalentProperties( $Q_1 \dots Q_n$ )	+	+
Individual( <i>o</i> type( $C_1$ ) ... type( $C_n$ ))	+	+
value( $R_1 o_1$ ) ... value( $R_n o_n$ )	+	+
SameIndividual( $o_1 \dots o_n$ )	+	-
DifferentIndividuals( $o_1 \dots o_n$ )	+	-*
<i>Descriptions</i>		
<i>A</i> (URI Reference)	+	+
owl:Thing	+	+
owl:Nothing	+	+
intersectionOf( $C_1 \dots C_n$ )	+	+
unionOf( $C_1 \dots C_n$ )	+	lhs
complementOf( <i>C</i> )	+	-
oneOf( $o_1 \dots o_n$ )	+	lhs
restriction( <i>Q</i> someValuesFrom ...)	+	lhs
restriction( <i>Q</i> allValuesFrom ...)	+	rhs
restriction( <i>Q</i> value ...)	+	+
restriction( <i>Q</i> minCardinality( <i>n</i> ))	+	-
restriction( <i>Q</i> maxCardinality( <i>n</i> ))	+	rhs
constraint( <i>R</i> someValuesFrom( <i>C</i> ))	-	rhs
constraint( <i>R</i> allValuesFrom( <i>C</i> ))	-	rhs
constraint( <i>R</i> value( <i>o</i> ))	-	rhs
constraint( <i>R</i> minCardinality( <i>n</i> ))	-	rhs
constraint( <i>R</i> maxCardinality( <i>n</i> ))	-	rhs
<i>Datatype Expressions and n-ary Datatype constraints</i>		
DatatypeExpression ( <i>E/n</i> )		
and/or/domain( $E_1/n \dots E_m/n$ )	-	+
restriction( $U_1 \dots U_n$ someValuesFrom( <i>E/n</i> ))	-	lhs
restriction( $U_1 \dots U_n$ allValuesFrom( <i>E/n</i> ))	-	rhs
* would be superfluous because OWL Flight adopts UNA		

Table 3: Features of OWL DL vs. OWL Flight

molecules can be combined, for instance  $r1 : Room[hasBooked \rightarrow 2]$ , stating that  $r1$  is an instance of the concept Room, having the value 2 for attribute *hasBooked*. Over such molecules we define clauses of the form

$$m \leftarrow m_1, \dots, m_n$$

with the usual meaning from logic programming, where the body molecules of such rule may be

preceded by the negation as failure symbol `not`. Furthermore we allow modules of the form  $x \neq y$ , representing inequality, in the body of such clauses. Variables in F-Logic might be used for any kind of objects, individuals, concepts or even attributes, allowing meta-modeling with sets of such clauses (i.e. F-Logic Programs). There exist efficient engines, such as Ontobroker [Decker et al., 1999] and FORA-2 [Yang et al., 2003], for evaluating these kinds of programs. Both implementations use a translation to plain Datalog in order to evaluate the program using a plain Datalog engine.

We will illustrate the mapping from OWL Flight to F-Logic with some small examples. Simple subclass axioms or partial class definitions for named classes, i.e. statements of the form `Class(A partial C1 ... Cn)` are translated to sets of statements of the form:

$$x : C_i \leftarrow x : A$$

Nested property restrictions in  $C_i$  can be handled by chaining variables, e.g.

```
Class(A partial restriction(R allvaluesFrom
  (restriction S allValuesFrom B)))
```

is translated to:

$$z : B \leftarrow x : A, x[R \rightarrow y], y[S \rightarrow z]$$

Particularly, constraints are translated to integrity constraints, i.e. clauses with an empty head, and involve negation as failure for checking integrity on the knowledge base. So, e.g.

```
Class(A partial constraint(R allValuesFrom(C)))
```

will result in

$$\leftarrow x : A[R \rightarrow y], \text{not } y : C$$

The complete translation [de Bruijn et al., 2004c] involves some particularities, but indeed allows us to translate all of OWL Flight to function-free F-Logic programs with integrity constraints and default negation, i.e. the Datalog<sup>IC,≠,not</sup> fragment of F-Logic. The Datatype-Expression facilities of OWL-E fit nicely in this translation.

Summarizing, OWL Flight on the one hand uses a maximal subset of OWL DL which is compatible with the logic programming world and on the other hand defines slight extensions addressing some of the limitations identified in Section 4. Although this new OWL “dialect” can not to be viewed as a fully-fledged ontology language to solve all the above mentioned problems, we conceive it as a solid starting point for a unified framework of OWL based ontology languages combining the benefits of the Description Logics and Logic Programming paradigms. Furthermore, it serves as a good starting point to investigate the limitations of OWL and ways to overcome its limitations.

## 6 Conceptual Modeling on the Semantic Web

In this section we compare the modeling constructs of OWL Flight and OWL DL in the context of modeling tasks on the Semantic Web. We discuss the major differences in modeling constructs between OWL DL and OWL Flight and discuss their merits on the Semantic Web.

We only discuss the modeling construct which differ significantly between OWL DL and OWL Flight. We do not discuss a number features of OWL DL which are lacking from OWL Flight, namely disjunction and classical negation.

## 6.1 Cardinality Restrictions and Constraints

OWL DL has the notion of *cardinality* restrictions, which can be used in class definitions. Cardinality restrictions allow for inferring equality and/or the existence of individuals not in the knowledge (see also Section 2. OWL Flight has an explicit notion of cardinality constraints, which are used to *check* the number of values for a certain property, rather than to derive equality or assume existence of individuals beyond the knowledge base.

Besides the possible non-intuitiveness of deriving equality (see also Section 4.3), we see the following pitfall for a language which allows to infer equality: The possibility to derive equality might be misleading, because not all equality on the Semantic Web can be resolved in the logical language.

Equality reasoning in OWL DL is not powerful enough to resolve all equalities between identifiers on the Semantic Web. Only if the world would be perfectly and completely modeled in an ontology and only if all individuals on the Semantic Web are related to this ontology could all equalities on the Semantic Web be resolved. We argue that very few equalities can actually be resolved with reasoning and that many derived equalities are actually faulty. Thus, it makes more sense in our opinion to either resolve equalities beyond the logical language or to make strong assumptions on the available knowledge, i.e. assume that each identifier in the knowledge base uniquely identifies an individual (the *unique name assumption*<sup>4</sup>).

## 6.2 Value restrictions vs. constraints

OWL allows for two kinds of value restrictions, namely existential and universal. However, an existential value restriction merely corresponds to a qualified minimal cardinality of 1. Therefore, and because we have already discussed cardinality restrictions in the previous section, we only treat universal value restrictions here.

When used in a class definitions, a value restriction over an object property can be used to derive additional information about property values with members of this class as its domain (see also Section 4.3). Take for example the following definition:

```
Class(Parent partial
  restriction(hasChild allValuesFrom Person))
```

For each individual which is a member of the class **Parent**, we can infer that each child is a member of the class **Person**. Notice that this axiom is not used to *check* whether the child is a person, but rather to *infer* the fact that each child is a person. Notice that if someone mistakenly asserts an individual which is not a person to be value for this property, the (incorrect) inference is made that this individual is a person.

OWL Flight allows for the specification of value *constraints*. A constraint is used to *check* whether the knowledge in the knowledge base conforms to the constraint. Take for example the following definition:

```
Class(Parent partial
  constraint(hasChild allValuesFrom Person))
```

This constraint does not allow any child which is not known to be a person.

---

<sup>4</sup>Notice that the unique name assumption can be introduced for any Description Logic knowledge base. However, all individuals in the knowledge base need to be enumerated.

We argue that the constraining approach is more intuitive to many computing professionals with Software Engineering and Database Systems backgrounds [de Bruijn et al., 2004b]. However, constraints do not allow additional inferences. Therefore, there is a tradeoff between the ability to detect modeling mistakes and the inferencing power of the language. OWL DL focuses more on the inferencing power of the language, whereas OWL Flight focuses more on catching modeling mistakes.

We have pointed out in Section 4.3 that in OWL DL there is an asymmetry between the treatment of datatype and abstract individuals. From a modeling point of view, this asymmetry is most apparent in the way property restrictions and range restrictions of datatype properties are handled, as we have argued in Section 4.3. Consider the following OWL DL knowledge base:

```
Class(A partial restriction(Q allValuesFrom D))
Individual(a value(Q b))
```

Say,  $Q$  is an object property and  $D$  is an abstract class description. By the universal value restriction, we can infer that  $b$  is an instance of  $D$ .

Now suppose  $Q$  is a datatype property and  $D$  is a datatype. From this knowledge base, we cannot conclude that  $b$  is a data value of type  $D$ , and thus the knowledge base is inconsistent.

This asymmetry between the treatment of object and datatype properties does not occur in OWL Flight, because it allows to model constraints for object properties which are treated in the same way as constraints for datatype properties.

### 6.3 Meta-modeling

Meta-modeling is based on the principle that the same object can be seen as a class, an individual or a property, depending on the point of view. An example taken from [Schreiber, 2002] is the object ‘Boeing747’ which is an instance of the class ‘AircraftType’ but is itself also a class whose instances are the individual aircraft. Another example [Noy et al., 2004] is the description of the topics of individual books. Each topic is a class and each book is an individual. This obviates the need to use classes as property values. Because of the heterogeneity to be expected on the Semantic Web, a conceptual modeling language for the Semantic Web should support meta-modeling.

RDFS and OWL Full allow full meta-modeling. Each identifier can denote a class, instance and/or property. Unfortunately, RDFS is not expressive enough for many applications and OWL Full is undecidable in general. The decidable species of OWL, namely OWL Lite and OWL DL, unfortunately do not allow meta-modeling. Instead, they require a strict separation between identifiers of classes, individuals and properties.

In the case of a strict separation between classes and instances, it is sometimes hard to find an agreement as to whether to model an object as a class or an instance. Since an ontology is shared by its very nature, it is important to find this agreement. In case inter-operation is required between different ontologies, this requirement becomes even more apparent, because the modelers of the different ontologies might have different notions of what is an instance and what is a class.

OWL Flight allows meta-modeling while retaining decidability, following the meta-modeling support of F-Logic [Kifer et al., 1995], which allows a limited form of meta-modeling by interpreting an identifier differently depending on the context in which it occurs. In this way, F-Logic stays inside the expressiveness of first-order logic. OWL Flight allows meta-modeling in the same way as F-Logic and by doing this, it stays inside the (decidable) Datalog fragment.

## 6.4 Complete Class Definitions

OWL allows two types of class definitions: partial class definitions and complete class definitions. A partial class definition corresponds with a necessary definition, thus, it specifies all conditions (such as superclasses and property restrictions) which are necessarily fulfilled (and thus *inferred*) for all members of the class. A complete class definition specifies necessary and sufficient conditions, which means that not only are these conditions inferred from class membership, but class membership is also inferred if all conditions are fulfilled, i.e. if an individual is a member of all superclasses and all property restrictions are fulfilled.

Complete class definitions are allowed only to a limited extent in OWL Flight: only named classes and individual value restrictions are allowed in the definition, because OWL Flight is based on Datalog, which only allows a single atom in the head of a rule and does not allow existential quantification (see also Table 3). In contrast, OWL DL allows all descriptions in both partial and complete class definitions.

To harvest the full power of Description Logic reasoning, complete class definitions should be used, because they allow for more powerful inference than partial definitions. However, it is hard to model complete definitions correctly, especially because it is easy to miss certain aspects of a class when creating the definition, in which case the definition is incorrectly labeled as ‘complete’. Thus, there is again a tradeoff between the ability to detect modeling mistakes and the inferencing power of the language.

**Example 6.1** *We can extend the OWL DL knowledge base of Example 4.1 with the following class definition:*

```
Class(FlightSeat complete Seat
restriction(hasPassenger maxCardinality(1))
restriction(hasFlight cardinality(1)))
```

*This is a necessary and sufficient definition of the concept *FlightSeat*, which means that every instance of *FlightSeat* is a *Seat*, has at most one passenger, and has exactly one flight associated with it, but also means that every instance of *Seat* which has at most one passenger and has exactly one flight associated with it is also an instance of *FlightSeat*.*

*Consider now the concept *AirportTaxiSeat*. This concept covers seats in a taxi of a special service which transports passengers to the airport. In order to identify at what time people should be at the airport, each seat is associated with a specific flight. We obtain the following partial class definition*

```
Class(AirportTaxiSeat partial Seat
restriction(hasPassenger maxCardinality(1))
restriction(hasFlight cardinality(1))
restriction(hasTaxi cardinality(1)))
```

*From these definitions we can infer that every instance of *AirportTaxiSeat* is also an instance of *FlightSeat*. This was most likely not the intention of the modeler.*

*In order to specify that the concepts *AirportTaxiSeat* and *FlightSeat* are disjoint we can add the following axiom to the knowledge base:*

```
DisjointClasses(AirportTaxiSeat FlightTaxi)
```

However, this would mean that there can be no instances of *AirportTaxiSeat*, since every instance of *AirportTaxiSeat* is necessarily an instance of *FlightSeat*, because of the complete definition of *FlightSeat*.

As we can see from the example, it can easily happen that classes are unintentionally related with each other in the case of complete class definitions. In case of partial definitions, it is always possible to express disjointness through explicit axioms. However, as we can see from the example, asserting disjointness in the presence of complete class definitions can lead to unsatisfiable concepts (i.e. concepts with an empty extension).

Correct and complete modeling cannot be guaranteed in general in such an open and distributed environment as the Web. However, in closed and controlled sections of the Semantic Web where it is possible to guarantee a certain degree of correctness of modeling, complete class definitions can be very useful. Complete class definitions are very much related with the subsumption reasoning task on which we will elaborate further in Section 7.1.

## 6.5 Negation

OWL DL allows for classical negation through the `complementOf` construct. “Classical” means that negation is the same as in (classical) first-order logic. First-order logic is monotonic, which means that conclusions drawn with current knowledge can not be invalidated when adding additional knowledge. A negated description  $\neg C$  is only true if it is *known* that  $C$  is false.

OWL Flight implements so-called *default negation* inside the minimal cardinality and the value constraints. The behavior of default negation is non-monotonic, which means that conclusions drawn with the current knowledge can be invalidated by introducing additional knowledge. A default negated description *not C* is true if  $C$  is *not known* to be false.

There are cases in which default negation is the most useful and cases in which classical negation is most useful. Consider the two sentences: (a) “if I do not know that a seat on a flight is reserved, I can reserve it” and (b) “if I know that a seat on a flight is not reserved, I can reserve it”. The sentence (a) is naturally expressed using default negation, whereas the sentence (b) is naturally expressed using classical negation. Default negation is typically used together with the closed-world assumption, which means that the knowledge in the knowledge base is complete, i.e. that all knowledge in the knowledge base is true and all knowledge not in the knowledge base is false.

Classical negation is very cautious in the sense that if you are not really sure that something is not true, then it does not allow to infer anything. In contrast, default negation allows to infer negative information from the absence of information and thus allows for many more inferences. However, the danger is that these inferences might be incorrect. Many practical applications, however, need a way to check for the absence of information to make certain inferences. For example, in order to check whether a seat can be booked, it is necessary to check whether the seat has been booked before. However, if there is no explicit information that the seat has not been booked, it will never be possible to book the seat.

## 7 Reasoning tasks on the Semantic Web

In this section we describe the primary reasoning tasks on the Semantic Web and how there are supported by OWL DL and OWL Flight. The primary reasoning tasks we consider are subsumption and query answering.

## 7.1 Subsumption

Subsumption corresponds with checking whether a class description subsumes (is more general than) another class description. By doing this for all classes in the knowledge base, one can compute the subsumption hierarchy. By checking the place in the subsumption hierarchy of a given class description, this class description is classified with respect to the knowledge base and hidden relationships with other classes in the knowledge base become visible.

Subsumption can be reduced to checking (un)satisfiability and thus, for OWL DL, an efficient Description Logic algorithm [Horrocks et al., 2000] can be used to compute the subsumption for any two OWL DL descriptions. There exist several implementations for the subset of OWL DL which excludes nominals, such as RACER [Haarslev and Möller, 2003] and FaCT++ [Tsarkov and Horrocks, 2004].

For OWL Flight, subsumption reasoning can be performed using a Description Logic reasoner for that subset of OWL Flight which falls inside OWL DL. There are also techniques for subsumption reasoning using Logic Programming engines, again using the same fragment of OWL Flight. For checking subsumption of named classes one can simply introduce a new individual, assert membership of the subsumed class and query for membership of the subsuming class [Grosz et al., 2003]. In order to check subsumption for complex descriptions, a new rule need to be added to the knowledge base which considerably slows down the reasoning process. These techniques [Grosz et al., 2003, Volz, 2004] effectively reduce subsumption reasoning to query answering in a deductive database. However, Description Logic reasoners are in general more efficient for subsumption reasoning (cf. [Volz, 2004]).

It is not entirely clear how constraints in OWL Flight interact with subsumption reasoning. Presumably, they should not be taken into account when doing subsumption reasoning, because constraints are not part of the logical theory (see also Section 2). This is a matter which requires further investigation.

The question that arises is: what are the use cases for classification and do we actually need it?

Particular use cases could be the classification of user goals with respect to web service descriptions in the field of Semantic Web Services for indexing goal and web service descriptions [Keller et al., 2004]. Another scenario is the diagnosis in the medical domain where a collection of observations could be classified with respect to an ontology of diagnoses to check which diagnosis fits with the observations.

Notice however that in case knowledge is modeled incorrectly or not to a sufficient degree, automatic classification can lead to unexpected inferences, which are correct with respect to the ontology, but not with respect to the real world, as was demonstrated in Section 6.4.

Subsumption reasoning allows inferring relationships between terms which have not been explicitly stated. On the one hand, this kind of reasoning might not be easily accepted in some areas. We believe that businesses will want control over their business vocabularies and do not want (possibly incorrect) inferred relationships between terms.

On the other hand, some areas may greatly benefit from the additional knowledge derived in subsumption reasoning. For example, search results in Knowledge Management applications might be improved because more relevant terms are taken into account and also other applications where 100% precision is not required or where complete and correct modeling can be assumed and where mistakes introduced through inferences can be tolerated can benefit greatly from classification reasoning. A domain where 100% precision is not required is Web Service indexing, because there

is always a human in the loop who verifies the results of the search. However, if this were to be combined with automated execution, precision is required.

There is a tradeoff between the possibilities for automation on the one hand and the requirements on the completeness and correctness of the modeling on the other hand. If one wants to automate certain tasks on the Semantic Web, such as Web Service execution, then correct and complete modeling of goals and web services is required. If one cannot assume correct and complete modeling of the functionality of a web service, then it is impossible to automate the execution.

## 7.2 Query Answering

There exist two different types of queries, namely ground queries and open queries. A ground query consists of a knowledge base and a ground fact. The inference task is to check whether the ground fact is entailed by the knowledge base. An open query is a formula with free variables. The query answer consists of a number of substitutions for the variables with values from the knowledge base. An open query can be reduced to a number of ground queries, namely, a ground query for each of the facts in the knowledge base. This is clearly not efficient and thus optimizations have been (and are being) developed for answering such open queries.

As we can see from Section 7.1, the main power of Description Logics lies in subsumption reasoning. Query answering reasoning is also possible, but is in general less investigated and very few optimizations exist to do query answering efficiently in Description Logics, as we have argued in Section 4.2.

In fact, the range of possibilities for ABox are just starting to be explored. The “classical” approach to ABox reasoning is for each individual  $a$  to assert a new class  $A$  in the TBox, which is subsumed by the class of which  $a$  is an instance. By classifying the TBox, the class membership of  $a$  consists of the classes which subsume  $A$ , thus query answering is reduced to classification. Other approaches are under development, such as the use of a database to contain the information about the instances and also part of the materialization of the ABox [Horrocks et al., 2004a], as well as indexing the ABox and caching query results [Haarslev and Möller, 2004]. Other approaches translate (part of) the Description Logic ontology to a (disjunctive) deductive database (e.g. [Grosz et al., 2003, Hustadt et al., 2004]).

OWL Flight was developed with efficient query answering using deductive databases in mind. Thus, common off-the-shelf deductive database engines (e.g. OntoBroker [Decker et al., 1999], FLORA-2 [Yang et al., 2003]) can be used for reasoning with OWL Flight. These deductive databases incorporate optimizations which have been developed in the database research area.

Summarizing, we can say that the expressive power, the modeling constructs and the current state-of-the-art reasoners for Description Logics (and thus OWL DL) are tuned to powerful subsumption reasoning. The expressive power of OWL Flight takes into account current state-of-the-art deductive databases and thus allows to take advantage of efficient query answering. OWL DL has significantly more expressive power than OWL Flight, because of disjunction and existential quantification allowed in the language. Therefore, even if optimizations for query answering with Description Logics are developed, these implementations cannot get around the theoretical complexity of the reasoning task.

	Conceptual Modeling	Reasoning tasks
OWL DL	<ul style="list-style-type: none"> <li>• Restrictions allow inferring new knowledge</li> <li>• Modeling style of Knowledge Representation</li> <li>• Complete class descriptions</li> </ul>	<ul style="list-style-type: none"> <li>• Efficient subsumption for a significant part (SHIQ)</li> <li>• Query Answering under-developed</li> </ul>
OWL Flight	<ul style="list-style-type: none"> <li>• Constraints check existing knowledge</li> <li>• Modeling style of Databases and Software Engineering</li> <li>• Limited complete class descriptions</li> </ul>	<ul style="list-style-type: none"> <li>• Limited expressiveness for classification; not optimized</li> <li>• Query Answering well-developed; many well-known optimization techniques and implementations</li> </ul>

Figure 1: OWL DL and OWL Flight: conceptual modeling and reasoning

## 8 Conclusions

In this paper we have introduced OWL Flight, a variant of OWL which is based on Logic Programming rather than Description Logics. We took Description Logic Programs [Grosz et al., 2003] (the intersection of Description Logics and Logic Programming) as a basis and extended it with cardinality and value constraints, meta-modeling and powerful datatype support (based on OWL-E [Pan and Horrocks, 2004]).

We have compared the modeling styles for OWL DL and OWL Flight, as well as the reasoning tasks supported by both languages. We summarize the modeling styles and reasoning tasks for both languages in Figure 1.

With respect to the modeling styles we can conclude that OWL DL has an inferring modeling style. Restrictions on property values and cardinalities are used by the reasoner to *infer* new knowledge from the existing knowledge, such as equality of individuals and class membership. OWL Flight has a constraining modeling style; it allows expressing cardinality and values constraints on properties. Such constraints are used by the reasoner to *check* whether the knowledge in the knowledge base corresponds with the conditions expressed in the constraints.

With respect to the supported reasoning tasks, we can conclude that OWL DL is most suited for the subsumption task and OWL Flight is most suited for the query answering task, because of the modeling styles of the languages and because of the optimized algorithms and implementations which have been developed for the respective modeling tasks.

We have also seen that in order to harvest the full power of OWL DL with respect to the classification reasoning task, there are certain requirements on the correctness and completeness of modeling. Therefore, it is most suited in restricted application domains. There exists a tradeoff between the inferencing power of a language and the ability to detect modeling mistakes.

We can conclude that different tasks on the Semantic Web need different styles of modeling and different types of reasoning. Therefore, we argue that it is not sufficient to restrict the Semantic Web to only one style of modeling for ontologies, which is currently done by OWL, in the form of Description Logics. Many applications on the Semantic Web would benefit from an ontology language which favors a constraining modeling style, such as OWL Flight. Furthermore, such a language might appeal more to software engineers and database designers, so that they more easily adopt the Semantic Web.

Interoperation between Description Logic-based and Logic Programming-based languages can be achieved either through their common subset [Grosz et al., 2003], through an interface between the languages [Eiter et al., 2004] or through a common superset. Such a common superset is not easy to define, because the default negation common in logic programming does not fit nicely in a

first-order framework. Nonmonotonic extensions such as the **K** operator [Donini et al., 1998] could help bridge the gap.

We argue that the Semantic Web requires a unifying framework for the Logic Programming and the Description Logic paradigm. Such a unifying framework would consist of a core language, based on [Grosz et al., 2003], which is the common subset of both paradigms. The framework should have extensions in both the Logic Programming and the Description Logic directions to allow applications on the Semantic Web to take advantage of existing algorithms and implementations. The languages would then be unified in a language which captures both paradigms. This unifying language could be a first-order language with nonmonotonic extensions in order to capture the nonmonotonicity of negation in Logic Programming. We are currently in the process of developing such a framework in the context of the WSML Working Group [de Bruijn et al., 2004a].

## Acknowledgement

We would like to thank Ian Horrocks, Boris Motik, Michael Kifer, and all members of the WSML working group for useful comments and discussions around the material presented in this paper.

## References

- [Baader et al., 2003] Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., and Patel-Schneider, P. F., editors (2003). *The Description Logic Handbook*. Cambridge University Press.
- [Berners-Lee et al., 2001] Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American*, 284(5):34–43.
- [Borgida, 1996] Borgida, A. (1996). On the relative expressiveness of description logics and predicate logics. *Artificial Intelligence*, 82(1–2):353–367.
- [Brickley and Guha, 2004] Brickley, D. and Guha, R. V. (2004). RDF vocabulary description language 1.0: RDF schema. Recommendation 10 February 2004, W3C. Available from <http://www.w3.org/TR/rdf-schema/>.
- [de Bruijn et al., 2004a] de Bruijn, J., Lausen, H., and Fensel, D. (2004a). The WSML Family of Representation Languages. Deliverable D16.1v0.2, WSML, <http://www.wsmo.org/wsml/>. Available from <http://www.wsmo.org/2004/d16/d16.1/v0.2/>.
- [de Bruijn et al., 2004b] de Bruijn, J., Polleres, A., Lara, R., and Fensel, D. (2004b). OWL<sup>-</sup>. Deliverable D20.1v0.2, WSML. Available from <http://www.wsmo.org/2004/d20/d20.1/v0.2/>.
- [de Bruijn et al., 2004c] de Bruijn, J., Polleres, A., Lara, R., and Fensel, D. (2004c). OWL flight. Deliverable D20.3v0.1, WSML. Available from <http://www.wsmo.org/2004/d20/d20.3/v0.1/>.
- [Dean and Schreiber, 2004] Dean, M. and Schreiber, G., editors (2004). *OWL Web Ontology Language Reference*. W3C Recommendation 10 February 2004.

- [Decker et al., 1999] Decker, S., Erdmann, M., Fensel, D., and Studer, R. (1999). *Ontobroker: Ontology based Access to Distributed and Semi-Structured Information*. Kluwer Academic.
- [Donini et al., 1998] Donini, F. M., Lenzerini, M., Nardi, D., Nutt, W., and Schaerf, A. (1998). An epistemic operator for description logics. *Artificial Intelligence*, 100(1–2):225–274.
- [Eiter et al., 2004] Eiter, T., Lukasiewicz, T., Schindlauer, R., and Tompits, H. (2004). Combining answer set programming with description logics for the semantic web. In *Proc. of the Int. Conf. of Knowledge Representation and Reasoning (KR04)*.
- [Fensel, 2003] Fensel, D. (2003). *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce, 2nd edition*. Springer-Verlag, Berlin.
- [Grosz et al., 2003] Grosz, B. N., Horrocks, I., Volz, R., and Decker, S. (2003). Description logic programs: Combining logic programs with description logic. In *Proc. Intl. Conf. on the World Wide Web (WWW-2003)*, Budapest, Hungary.
- [Haarslev and Möller, 2003] Haarslev, V. and Möller, R. (2003). Racer system description. In *Proc. of the 12th Int. World Wide Web Conf. (WWW2003)*, pages 48–57. ACM.
- [Haarslev and Möller, 2004] Haarslev, V. and Möller, R. (2004). Optimization techniques for retrieving resources described in OWL/RDF documents: First results. In *Ninth Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR2004)*, pages 163–173, Whistler, BC, Canada.
- [Heflin and Muñoz-Avila, 2002] Heflin, J. and Muñoz-Avila, H. (2002). LCW-based agent planning for the semantic web. In *Ontologies and the Semantic Web. Papers from the 2002 AAAI Workshop WS-02-11*, pages 63–70, Menlo Park, CA, USA. AAAI Press.
- [Horrocks et al., 2004a] Horrocks, I., Li, L., Turi, D., and Bechhofer, S. (2004a). The instance store: DL reasoning with large numbers of individuals. In *Proc. of the 2004 Description Logic Workshop (DL 2004)*, pages 31–40.
- [Horrocks and Patel-Schneider, 2003] Horrocks, I. and Patel-Schneider, P. F. (2003). Reducing OWL entailment to description logic satisfiability. In *Proc. of the 2003 Int. Semantic Web Conf. (ISWC 2003)*, Sanibel Island, Florida.
- [Horrocks et al., 2004b] Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosz, B., and Dean, M. (2004b). SWRL: A semantic web rule language combining OWL and RuleML. Member submission 21 may 2004, W3C. Available from <http://www.w3.org/Submission/SWRL/>.
- [Horrocks et al., 2003] Horrocks, I., Patel-Schneider, P. F., and van Harmelen, F. (2003). From SHIQ and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26.
- [Horrocks and Sattler, 2001] Horrocks, I. and Sattler, U. (2001). Ontology reasoning in the SHOQ(D) description logic. In *Proc. of the Seventeenth Int. Joint Conf. on Artificial Intelligence (IJCAI2001)*.

- [Horrocks et al., 2000] Horrocks, I., Sattler, U., and Tobies, S. (2000). Practical reasoning for very expressive description logics. *Logic Journal of the IGPL*, 8(3):239–264.
- [Hustadt et al., 2004] Hustadt, U., Motik, B., and Sattler, U. (2004). Reducing SHIQ<sup>-</sup> description logic to disjunctive logic programs. In *Proc. of the 9th Int. Conf. on Knowledge Representation and Reasoning KR2004*, pages 152–162, Whistler, Canada.
- [Keller et al., 2004] Keller, U., Lara, R., and Polleres, A., editors (2004). *WSMO Discovery*. WSMO Working Draft D5.1v0.1. Available from <http://www.wsmo.org/2004/d5/d5.1/v0.1/>.
- [Kifer et al., 1995] Kifer, M., Lausen, G., and Wu, J. (1995). Logical foundations of object-oriented and frame-based languages. *JACM*, 42(4):741–843.
- [Levy and Rousset, 1998] Levy, A. Y. and Rousset, M.-C. (1998). Combining horn rules and description logics in CARIN. *Artificial Intelligence*, 104:165 – 209.
- [Noy et al., 2004] Noy, N., Hayes, P., McBride, B., Rector, A., and Vatant, B. (2004). Representing classes as property values on the semantic web. Working draft, W3C. Available from <http://www.w3.org/TR/swbp-classes-as-values>.
- [Pan and Horrocks, 2004] Pan, J. Z. and Horrocks, I. (2004). OWL-E: Extending OWL with expressive datatype expressions. IMG Technical Report IMG/2004/KR-SW-01/v1.0, Victoria University of Manchester. Available from <http://dl-web.man.ac.uk/Doc/IMGTR-OWL-E.pdf>.
- [Patel-Schneider et al., 2004] Patel-Schneider, P. F., Hayes, P., and Horrocks, I. (2004). OWL web ontology language semantics and abstract syntax. Recommendation 10 February 2004, W3C.
- [Ramakrishnan, 1991] Ramakrishnan, R. (1991). Magic templates: A spellbinding approach to logic programs. *Journal of Logic Programming*, 11(3-4):189–216.
- [Schreiber, 2002] Schreiber, G. (2002). The web is not well-formed. *IEEE Intelligent Systems*, 17(2). Contribution to the section Trends and Controversies: Ontologies KISSES in Standardization.
- [Tsarkov and Horrocks, 2004] Tsarkov, D. and Horrocks, I. (2004). Efficient reasoning with range and domain constraints. In *Proc. of the 2004 Description Logic Workshop (DL 2004)*, pages 41–50.
- [Volz, 2004] Volz, R. (2004). *Web Ontology Reasoning with Logic Databases*. PhD thesis, AIFB, Karlsruhe.
- [Yang et al., 2003] Yang, G., Kifer, M., and Zhao, C. (2003). FLORA-2: A rule-based knowledge representation and inference infrastructure for the semantic web. In *Proc. of the Second Int. Conf. on Ontologies, Databases and Applications of Semantics (ODBASE)*, Catania, Sicily, Italy.