Expressing DL-Lite Ontologies with Controlled English

Raffaella Bernardi, Diego Calvanese, Camilo Thorne

Faculty of Computer Science
Free University of Bozen-Bolzano
Piazza Domenicani 3, Bolzano, Italy
{bernardi,calvanese,cthorne}@inf.unibz.it

Abstract. We are interested in providing natural language front-ends to databases upon which an ontology layer has been added. Specifically, here we deal with how to express ontologies formalized in Description Logics in a controlled language, i.e., a fragment of natural language tailored to compositionally translate into a knowledge representation (KR) language. As KR language we have chosen *DL-Lite*$_{R,\sqcap}$, a representative of the well-known *DL-Lite* family [3, 4], and we aim at understanding the kind of English constructs the controlled language can and cannot have to correspond to *DL-Lite*$_{R,\sqcap}$. Hence, we compare the expressive power of *DL-Lite*$_{R,\sqcap}$ to that of various fragments of FOL identified by Pratt and Third as corresponding to fragments of English [8]. Our analysis shows that *DL-Lite*$_{R,\sqcap}$, though itself tractable, is incomparable in expressive power with respect to tractable fragments of English. Interestingly, it allows one to represent a restricted form of relative clauses, which lead to intractability when used without restrictions on the occurrences of negations, and existential quantifiers.

1 Introduction

The importance of using an ontology to facilitate the access of users to structured data is well established [2, 3]. Having an ontology as support for querying a database (DB) will allow the user to find the relevant answers without knowing about the structure of the DB itself. Though having an ontology will provide support to users able to use query languages, it will still fail to make the data accessible to non expert users. These could instead benefit from using a natural language interface to the ontology and the DB, both for querying the DB and for entering knowledge, either intensional (i.e., ontology assertions) or extensional (i.e., DB facts) one. Therefore, we are interested in looking at the *query entailment problem*, i.e., $\mathcal{T} \cup \mathcal{D} \models \varphi$, for an ontology $\mathcal{T}$, database $\mathcal{D}$, and query φ, but from a natural language perspective.

We know that query entailment can be done efficiently (i.e., in LOGSPACE in the size of the DB $\mathcal{D}$), if the ontology $\mathcal{T}$ is expressed in a Description Logic (DL) of the *DL-Lite* family [3, 4] and the query φ is a (union of) conjunctive queries (CQs). When resorting to natural language interfaces, we aim at preserving this efficiency. Thus, we are interested in understanding (*i*) which fragments of natural language correspond to the two fragments of First-Order Logic (FOL) we need, viz. *DL-Lite* and CQs, and (*ii*) whether these two fragments will be suitable for non expert users to accomplish the tasks we are interested in, entering intensional and extensional knowledge into an ontology and querying a DB.

Roughly, with respect to FOL, CQs lack negation and universal quantification. This might seem too restrictive when interested in expressing natural language questions as DB queries. However, an analysis of several corpora of real life users' questions[1] has shown that the use of those operators in questions is rather limited. Similarly, we are now trying to understand how far *DL-Lite* is from the linguistic structures that domain experts would naturally use to describe their intensional knowledge. To this end, as a preliminary study, we have started looking at the answers provided by domain experts to FAQs[2]. Again, the first results are rather promising, showing that domain experts, when allowed to freely use natural language, write rather simple structures with only few occurrences of those operators "forbidden" by *DL-Lite* definition, e.g., universal quantifiers in non subject position. Similarly, the use of negation and disjunction is rather limited and controlled while relative pronouns instead are rather common in these corpora and they are usually used to further specify properties of the nearest noun. As will become clearer in the next section, these operators are relevant to understand the connection between *DL-Lite* and natural language fragments since their corresponding logical operators are the major players in determining the complexity property of the entailment problem above.[3]

Against this background, our research line is as follows. We propose to study the problem of accessing structured data via an ontology by moving back and forth between logic and natural language: on the one hand, by studying the expressivity of suitable logic fragments and identifying the corresponding natural language fragments, and on the other hand, by analysing natural language structures used in real life applications and trying to extend the corresponding logic fragments to better suit users' needs, but without paying in terms of computational complexity.

In this paper, we concentrate on *DL-Lite*$_{R,\sqcap}$, which is the DL that stays as close as possible to the expressive power required to capture natural language constructs, while still preserving the nice computational properties of the *DL-Lite* family. As a first step towards understanding the relationship between ontology languages and natural language constructs, we compare *DL-Lite*$_{R,\sqcap}$ with the expressive power of fragments of FOL studied by Pratt and Third [8, 10] and defined starting from natural language.

2 The Description Logic *DL-Lite*$_{R,\sqcap}$

In this work, we consider a DL belonging to the *DL-Lite* family [3, 4], and specifically, we consider *DL-Lite*$_{R,\sqcap}$, in which the TBox is constituted by a set of (concept and role) *inclusion assertions* of the form $Cl \sqsubseteq Cr$ and $R_1 \sqsubseteq R_2$, where Cl and Cr denote concepts that may occur respectively on the left and right-hand side of inclusion assertions, and R_1, R_2 denote roles, constructed according to the following syntax:

$$\begin{array}{lcl@{\qquad\qquad}lcl} Cl & \longrightarrow & A \mid \exists R \mid Cl_1 \sqcap Cl_2 & R & \longrightarrow & P \mid P^- \\ Cr & \longrightarrow & A \mid \exists R \mid Cr_1 \sqcap Cr_2 \mid \exists R.A \mid \neg A \mid \neg \exists R & & & \end{array}$$

[1] http://wiki.answers.com/Q/WikiFAQs:Finding_Questions_to_Answer
http://clinques.nlm.nih.gov/JitSearch.html (clinical questions)

[2] http://www.unibz.it/library/faq/

[3] What seems to cause a lack of expressivity is the limitation on the occurrences of qualified existential, viz., the fact that they cannot occur in the left concepts of TBox statements, in all of the logics of the *DL-Lite* family. This will be the topic of further studies.

where A denotes an atomic concept, and P denotes an atomic role.

For convenience w.r.t. what we need in the following sections, we formally specify the semantics of *DL-Lite*$_{R,\sqcap}$, by providing its translation to FOL. Specifically, we map each concept C (we use C to denote an arbitrary concept, constructed applying the rules above) to a FOL formula $\varphi(C,x)$ with one free variable x (i.e., a unary predicate), and each role R to a binary predicate $\varphi(R,x,y)$ as follows:

$$\begin{array}{ll}
\varphi(A,x) = A(x) & \varphi(\exists R,x) = \exists y(\varphi(R,x,y)) \\
\varphi(\neg C,x) = \neg\varphi(C,x) & \varphi(\exists R.C,x) = \exists y(\varphi(R,x,y) \land \varphi(C,y)) \\
\varphi(C_1 \sqcap C_2,x) = \varphi(C_1,x) \land \varphi(C_2,x) & \\
\varphi(P,x,y) = P(x,y) & \varphi(P^-,x,y) = P(y,x)
\end{array}$$

Inclusion assertions $Cl \sqsubseteq Cr$ and $R_1 \sqsubseteq R_2$ of the TBox correspond then, respectively, to the universally quantified FOL sentences:

$$\forall x(\varphi(Cl,x) \rightarrow \varphi(Cr,x)) \qquad\qquad \forall x\forall y(\varphi(R_1,x,y) \rightarrow \varphi(R_2,x,y))$$

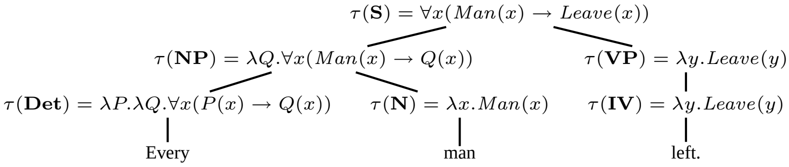
In *DL-Lite*$_{R,\sqcap}$, an ABox is constituted by a set of assertions on *individuals*, of the form $A(c)$ or $P(a,b)$, where A and P denote respectively an atomic concept and an atomic role, and a, and b denote constants. As in FOL, each constant is interpreted as an element of the interpretation domain, and we assume that distinct constants are interpreted as distinct individuals, i.e., we adopt the *unique name assumption* (UNA). However, in *DL-Lite*$_{R,\sqcap}$, we may drop such an assumption without affecting the complexity of reasoning, as established below. The above ABox assertions correspond to the analogous FOL facts, or, by resorting to the above mapping, to $\varphi(A,x)(c)$ and $\varphi(R,x,y)(a,b)$, respectively.

The reasoning services of interest for *DL-Lite*$_{R,\sqcap}$ knowledge bases are the standard ones, namely *knowledge base satisfiability*, and concept and role *satisfiability*, and *subsumption*. It has been shown in [4] that in *DL-Lite*$_{R,\sqcap}$ all such reasoning services are polynomial in the size of the knowledge base, and LOGSPACE in the size of the ABox only, i.e., in *data complexity*. Moreover, answering conjunctive queries whose atoms have as predicates atomic concepts and roles of a knowledge base, is also polynomial in the size of the knowledge base and in LOGSPACE in data complexity [3, 4].

3 Fragments of English

In this section we give a brief overview of Third and Pratt's controlled fragments of English (cf. [8]). They are subsets of standard English meant to capture some simple, albeit for our purpose important, structure of English. Their interest, as we said in the introduction, lies in the fact that we would like to know which subset of English we can use to express only those data constraints required by ontology-driven data access. It is thus crucial to know which natural language constructs express these constraints and, more specifically, those suitable for a *DL-Lite* ontology.

The key feature of these fragments of English is that they compositionally translate, modulo the standard semantic mapping foreseen by montagovian natural language formal semantics (cf. [6, 7]) into several fragments of FOL. Roughly: (*i*) A parse tree is computed. (*ii*) A FOL formula enriched with lambda operators from the lambda calculus is assigned to the words, i.e., the terminal nodes of the tree, representing their set-theoretical meaning. (*iii*) The logical formula representing the meaning of the parsed sentence is computed bottom-up by means of function application and beta-reduction at each internal node or component of the tree. This yields, ultimately, a FOL closed formula for the whole utterance called its *meaning representation* (MR). An example is given in the parse tree below, where τ returns the current value of the translation at each node.


$\tau(\mathbf{S}) = \forall x(Man(x) \rightarrow Leave(x))$
$\tau(\mathbf{NP}) = \lambda Q.\forall x(Man(x) \rightarrow Q(x))$
$\tau(\mathbf{VP}) = \lambda y.Leave(y)$
$\tau(\mathbf{Det}) = \lambda P.\lambda Q.\forall x(P(x) \rightarrow Q(x))$
$\tau(\mathbf{N}) = \lambda x.Man(x)$
$\tau(\mathbf{IV}) = \lambda y.Leave(y)$
Every
man
left.

For instance, by applying the translation procedure described above we get the following MRs from their corresponding English utterances:

1. Some people are weak. $\leadsto$ $\exists x(People(x) \wedge Weak(x))$.
2. Every husband has a wife. $\leadsto$ $\forall x(Hasband(x) \rightarrow \exists y(Wife(y) \wedge Has(x,y)))$.
3. Every salesman sells some merchandise to some customer. $\leadsto$ $\forall x(Salesman(x) \rightarrow \exists y(Customer(y) \wedge \exists z(Merchandise(z) \wedge Sells(x,z,y)))$.

Note that in 2. and 3. above, other translations might be possible due to NL ambiguity. However, these are discarded by the grammar studied by Third and Pratt that generates MRs following exclusively the surface order of components.

Schematically, the sentences above have the shape “Det N VP”, where the verb phrase (VP) is the constituent built out of a verb and its complements. We come back to this schema later to summarize the kind of constructs corresponding to *DL-Lite*$_{R,\sqcap}$.

The fragments of English themselves are built step by step, by starting with copula, nouns, negation, and the universal and existential quantifiers and by extending coverage to larger portions of English – covering relative constructions, ditransitive verbs, anaphora, as summarized by the table below.

This analysis is important for our purposes Because each NL construct has a meaning representation built out of some constant or some logical operation in FOL: The MRs of relatives (e.g., “who”) are built by conjunction ($\wedge$); negations (e.g., “no”, “not”) introduce logical negation ($\neg$); intransitive verbs (e.g., “runs”) and nouns (e.g., “man”) correspond to unary predicates; transitive verbs (e.g., “loves”) correspond to binary predicates, and ditransitive verbs (e.g., “sells to”) to ternary predicates; universal quantifiers (“every”, “all”, “everyone”) to $\forall$, and existential (“some”, “someone”) to $\exists$.

By building a family of fragments, Pratt and Third [8] have studied the impact on expressive power and computational complexity these constructs have (see Figure 1). As the reader can see, this process leads ultimately to an undecidable fragment of English. As a matter of fact, only the first two fragments, COP and COP+TV+DTV are tractable. Notice that as soon as we add rules dealing with the relative clause, we lose tractability. COP+Rel (i.e., COP with relative clauses), is already NP-Complete. COP+TV+DTV+Rel is NEXPTIME-Complete. This is because, as we said, relatives express conjunctions which, together with negation, generate logics (i.e., fragments of FOL) that contain the propositional calculus. But covering relatives to a certain extent is crucial: as we mentioned before, they occur quite frequently in NL utterances.

Fragment	**Coverage**	**Sat. decision class**
COP	Copula, common and proper nouns, negation, universal and existential quantifiers	P
COP+TV+DTV	COP + Transitive verbs (e.g. ”reads”) + Ditranstive verbs (e.g., ”sells”)	P
COP+Rel	COP + Relative pronoun (i.e., ”who”, ”that”, ”which”, etc.)	NP-Complete
COP+Rel+TV	COP + Transtive verbs + Relative pronoun	ExpTime-Complete
COP+Rel+TV+DTV	COP+TV+DTV + Relative pronouns	NExpTime-Complete
COP+Rel+TV+RA	COP+Rel+TV + Restricted anaphora	NExpTime-Complete
COP+Rel+TV+GA	COP+Rel+TV + Generalized anaphora	undecidable

Fig. 1. Fragments of English studied by Pratt and Third [8].

Some means to cover them without yielding an exponential blowup should be found. As shown in [1], this is possible if we choose as MR logic *DL-Lite*$_{R,\sqcap}$, which allows relatives ($\wedge$) to occur both in subject and in predicate position of sentences with an universal quantified subject, i.e., in the left and right concepts, respectively, of inclusion assertions.

COP and COP+TV+DTV generate, through this process of compositional translation described above, the following FOL fragments:

COP	**COP+TV+DTV**
$\pm A_1(c)$	ψ
$\exists x_1(A_1(x_1) \wedge \pm A_2(x_1))$	$Q_1 x_1(A_1(x_1) \boxdot \psi(x_1))$
$\forall x_1(A_1(x_1) \rightarrow \pm A_2(x_1))$	$Q_1 x_1(A_1(x_1) \boxdot \pm Q_2 x_2(A_2(x_2) \boxdot \psi(x_1, x_2)))$
	$Q_1 x_1(A_1(x_1) \boxdot \pm Q_2 x_2(A_2(x_2) \boxdot \pm Q_3 x_3(A_3(x_3) \boxdot \psi(x_1, x_2, x_3))))$

In the table above, $Q_i \in \{\forall, \exists\}$, for $1 \leq i \leq 3$, $\boxdot \in \{\wedge, \rightarrow\}$, c is an individual constant, the A_i's, for $1 \leq i \leq 3$, are unary predicates, and ψ is an n-place *literal* (postive or negative) over the variables $\{x_1, \ldots, x_n\}$ containing possibly constants (thus ψ is a grounded literal). A quick glance at these logic fragments tells us that they can express IS-A constraints, as well as ABoxes almost directly. Moreover, we can express unary and binary (and even ternary) predicates, together with quantification.

So the questions now are: exactly to what extent these two tractable fragments or, equivalently, the FOL fragments thereof generated can express *DL-Lite*$_{R,\sqcap}$? How much of *DL-Lite*$_{R,\sqcap}$ can not be expressed by these fragments? Answering these questions will shed light on the issue of which NL constructs can ultimately express *DL-Lite*$_{R,\sqcap}$ ontologies in a subset of English.

4 Comparing Expressive Power

In this section, we compare the expressive power of *DL-Lite*$_{R,\sqcap}$ with that of the two tractable fragments of English COP and COP+TV+DTV. We show that, under certain conditions, COP is contained in *DL-Lite*$_{R,\sqcap}$, as it should be expected, but that COP+TV+DTV only overlaps with *DL-Lite*$_{R,\sqcap}$. This is interesting, since, as shown in [1], Lite English, the controlled language that compositionally translates into *DL-Lite*$_{R,\sqcap}$, covers relative pronouns (mirrored by the qualified existential $\exists R.C$) without yielding an exponential blowup, as is the case with Pratt's fragments.

We begin by recalling some basic notions of FOL (without function symbols) model theory. An *interpretation structure* over a FOL signature (without function symbols) $\mathcal{L}$ is a tuple $\mathfrak{M} = \langle M; \{R_i^{\mathfrak{M}}\}_{i\in I}; \{c_j^{\mathfrak{M}}\}_{j\in J}\rangle$ where the $R_i^{\mathfrak{M}}$ are n-ary relations over M and the $c_j^{\mathfrak{M}}$ distinguished elements of M, for $i \in I, j \in J$. A structure $\mathfrak{M}'$ is said to be an *extension* of $\mathfrak{M}$ whenever the relations of $\mathfrak{M}$ are contained in those of $\mathfrak{M}'$ and they coincide on the distinguished elements. A structure is said to be a *model* of a sentence or formula ϕ whenever $\mathfrak{M} \models \phi$. The sentence ϕ *characterizes* the classes of its models (i.e., the class $\{\mathfrak{M} | \mathfrak{M} \models \phi\}$). These classes are called *properties*. The *expressive power* of a fragment of FOL is then formally given by the model theoretic properties its sentences can characterize. Finally, a FOL fragment Λ' is said to be *as expressive as* a fragment Λ when, and only when, Λ' can express all properties of Λ [9]. The idea of the proofs is to individuate properties expressible in one logic and not in the other – that is, classes of structures that *DL-Lite*$_{R,\sqcap}$ expresses but that COP+TV+DTV and COP may or may not express.

Theorem 1. *DL-Lite*$_{R,\sqcap}$ *is as expressive as COP, assuming the unique name assumption does not hold.*

Proof. Some COP sentences cannot be *a priori* expressed in *DL-Lite*$_{R,\sqcap}$. In particular, as we have seen, *DL-Lite*$_{R,\sqcap}$ as it is, cannot express negative facts: ABox assertions (i.e., ground atoms) cannot be negative following the standard definition of *DL-Lite*$_{R,\sqcap}$. However, we can easily express a negative fact $\neg A(c)$, by extending our signature with a new concept name A', and introducing the disjointness assertion $A' \sqsubseteq \neg A$ and the membership assertion $A'(c)$. To deal with COP formulas of the form $\exists x(P(x) \wedge Q(x))$ we proceed as follows: (*i*) we skolemise and extend our signature by adding a new constant c (expanding models $\mathfrak{M}$ to their skolem expansion $(\mathfrak{M}, c^{\mathfrak{M}})$) and (*ii*) we drop the unique name assumption (UNA) regarding constants when it comes to these new constants produced by skolemisation. We can then express these statements as ABox assertions $P(c)$ and $Q(c)$. □

Theorem 2. *DL-Lite*$_{R,\sqcap}$ *is not as expressive as COP+TV+DTV.*

Proof. To prove this result, we exhibit a closure property of *DL-Lite*$_{R,\sqcap}$ that is not preserved by COP+TV (and *a fortiori* by COP+TV+DTV). The formulas in *DL-Lite*$_{R,\sqcap}$ are all FOL $\forall\exists$ formulas, *modulo* the standard translation. A $\forall\exists$ formula or sentence is a formula $\phi := \forall x_1 \cdots \forall x_n \exists x_1 \cdots \exists x_m \psi$, for $n, m \geq 0$, where ψ is quantifier-free. Now, $\forall\exists$ formulas are closed under the *union of chains* property [5], defined as follows. We say that a formula ϕ is *closed under union of chains* iff for every partial order $\langle T, \prec_T\rangle$, for every model $\mathfrak{M}$ of ϕ, and every family $\{\mathfrak{M}_t\}_{t\in T}$ of extensions of $\mathfrak{M}$, s.t. $i \prec_T j$ implies that $\mathfrak{M}_j$ is an extension of $\mathfrak{M}_i$, for $i, j \in T$, then the structure $\mathfrak{M}_\omega$, called the *union structure* and defined below, is also a model of ϕ:

1. $M_\omega = \bigcup_{t\in T} M_t$
2. Every $R_i^{\mathfrak{M}_\omega}$ is the union of all the relations of the same arity and position among the $\mathfrak{M}_t$'s, for $t \in T$, $i \in I$
3. Every $c_j^{\mathfrak{M}_\omega}$ is a distinguished element among the the $\mathfrak{M}_t$'s, for $t \in T$, $j \in J$.

Therefore, every set of *DL-Lite*$_{R,\sqcap}$ sentences (assertions) will be closed under union of chains. Now, suppose, towards a contradiction that every property that is expressible in

COP+TV+DTV is expressible also in *DL-Lite*$_{R,\sqcap}$, in particular: $\exists x(P(x) \wedge \forall y(Q(y) \rightarrow R(x,y)))$. That is, after prenexing: $\exists x \forall y(P(x) \wedge (Q(y) \rightarrow R(x,y)))$. This sentence should be closed under union of chains, following the hypothesis. But this does not hold. To show this define a model $\mathfrak{M}$ of this sentence as follows: $M = \mathbb{N}$; $P^{\mathfrak{M}} = Q^{\mathfrak{M}} = M$; $R^{\mathfrak{M}} = \leq_{\mathbb{N}}$ (i.e., the usual loose order over positive integers).

Define next a sequence $\{\mathfrak{M}_i\}_{i \in \mathbb{N}}$ of extensions of $\mathfrak{M}$ as follows:

- $\mathfrak{M}_0$: $M_0 = M \cup \{e_0\}$; $P^{\mathfrak{M}_0} = Q^{\mathfrak{M}_0} = M_0$; $R^{\mathfrak{M}_0} = R^{\mathfrak{M}} \cup \{\langle e_0, 0\rangle\}$.
- $\mathfrak{M}_{i+1}$: $M_{i+1} = M_i \cup \{e_{i+1}\}$; $P^{\mathfrak{M}_{i+1}} = Q^{\mathfrak{M}_{i+1}} = M_{i+1}$; $R^{\mathfrak{M}_{i+1}} = R^{\mathfrak{M}_i} \cup \{\langle e_{i+1}, e_i\rangle\}$.

Now, $\{\mathfrak{M}_i\}_{i \in \mathbb{N}}$ constitutes a chain, since (*i*) a sequence is a family, (*ii*) $\langle \mathbb{N}, \leq_{\mathbb{N}}\rangle$ is a partial order and (*iii*) whenever $i \leq_{\mathbb{N}} j$, $\mathfrak{M}_j$ extends $\mathfrak{M}_i$. Finally, consider the union structure $\mathfrak{M}_\omega$ for this chain. $\mathfrak{M}_\omega$ is not a model of $\exists x \forall y(P(x) \wedge (Q(y) \rightarrow R(x,y)))$, since the relation $R^{\mathfrak{M}_\omega}$ of $\mathfrak{M}_\omega$ has no least element. □

Theorem 3. *COP+TV+DTV is not as expressive as DL-Lite*$_{R,\sqcap}$.

Proof. A *DL-Lite*$_{R,\sqcap}$ inclusion assertion of the form $\exists R \sqsubseteq A$ corresponds to the FOL sentence $\forall x \exists y(R(x,y) \rightarrow A(x))$. Skolemizing and clausifying this sentence yields: $\neg R(x, f(x)) \vee A(x)$, i.e., a clause containing both a positive unary literal and a binary negative literal containing function symbols. But in [8] it is proven that this particular kind of clauses lies beyond COP+TV+DTV, whence the result. □

Theorem 4. *COP and COP+TV+DTV overlap in expressive power with DL-Lite*$_{R,\sqcap}$.

Proof. Consider this following typical meaning representation formula for COP: $\forall x(P(x) \rightarrow Q(x))$. The models of these sentences are the FOL interpretation structures $\mathfrak{M} = \langle M; P^{\mathfrak{M}}, Q^{\mathfrak{M}}\rangle$, where $P^{\mathfrak{M}} \subseteq Q^{\mathfrak{M}}$. But this property can be easily expressed in the *DL-Lite*$_{R,\sqcap}$ with inclusion assertions. □

We finish by remarking that it can also be proved that COP is not as expressive as *DL-Lite*$_{R,\sqcap}$ either, since it cannot express binary relations. Moreover, we can show, in a way analogous to Theorem 2, that COP+TV is not as expressive as *DL-Lite*$_{R,\sqcap}$, by exhibiting a closure property of COP+TV, namely COP+TV-*simulation* [10] that is not verified by *DL-Lite*$_{R,\sqcap}$.

An understanding of how the different expressivity of the compared logics is reflected on the corresponding natural language fragments can be reached by considering the general schema “Det N VP” mentioned previously. In these fragments “Det” can be either “every” or “some” or “no”, and all of these determiners can build the direct and/or indirect objects of the transitive and ditransitive verbs, i.e., the complements in the “VP”; the verb of the “VP” can be negated. Logic conjunction is introduced only by the meaning representation of “some”. The sentences whose meaning representation belong to *DL-Lite*$_{R,\sqcap}$, instead, do not have ditransitive verbs (ternary relations) and in the “Det” position only the determiners “every” and “no” can occur. Notice, that the “N” and “VP” correspond respectively to the Cl and Cr concepts of a *DL-Lite*$_{R,\sqcap}$ inclusion assertion. As in the latter, the two parts can be complex: the “N” constituent can be a complex structure built out of a relative pronoun (e.g., “student who left”, “student who knows something"); transitive verbs can occur only with an unqualified existential as object; the verb of the relative clause (e.g., "left" and "knows", resp.) cannot be negated (negation does not occur in Cl), and the relative clause cannot be iterated, i.e., it cannot be used to modify the object of a transitive verb (only unqualified existential can occur in Cl). Similarly, the "VP" can be a complex structure: since it corresponds to the Cr concept, copula, intransitive verbs, and transitive verbs, with unqualified existential as object, can be negated. Whereas transitive verbs with qualified existential as object cannot be negated, e.g., "every student does not know something that is interesting" (notice how the relative clause modifies an existential building a qualified existential).

Finally, neither in COP and COP+TV+DTV nor in our fragment reflexive pronouns (e.g., "itself") and possessive (e.g., "their") are allowed: they would correspond to the introduction of role-value-maps, which is a notoriously problematic construct that may lead to undecidability.

5 Conclusions

We have compared the expressive power of *DL-Lite*$_{R,\sqcap}$ with that of Pratt's and Third's tractable fragments of English [8, 10]. Using model theoretic arguments, we have shown that the compared logics are incomparable to each other, even though a reasonable deal of the semantic structures captured by the two tractable fragments of English is shared by *DL-Lite*$_{R,\sqcap}$. We remark that a controlled natural language covering constructs such as intransitive and transitive verbs, copula, common nouns, adjectives, restricted occurrences of universal and existential quantification as well as of negation and relative pronouns can be "reverse engineered" from *DL-Lite*$_{R,\sqcap}$, as shown in [1].

Acknowledgements. Research partially supported by the FET project TONES (Thinking ONtologiES), funded by the EU under contract FP6-7603, and by the PRIN 2006 project NGS (New Generation Search), funded by MIUR.

References

1. R. Bernardi, D. Calvanese, and C. Thorne. Lite natural language. In *Proc. of the 7th Int. Workshop on Computational Semantics (IWCS-7)*, 2007.
2. A. Borgida, R. J. Brachman, D. L. McGuinness, and L. A. Resnick. CLASSIC: A structural data model for objects. In *Proc. of ACM SIGMOD*, pages 59–67, 1989.
3. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. DL-Lite: Tractable description logics for ontologies. In *Proc. of AAAI 2005*, pages 602–607, 2005.
4. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Data complexity of query answering in description logics. In *Proc. of KR 2006*, pages 260–270, 2006.
5. R. Cori and D. Lascar. *Logique mathématique (2 vols)*. Dunod, 2003.
6. L. T. F. Gamut. *Logic, Language and Meaning (2 vols.)*. University of Chicago Press, 1991.
7. R. Montague. The proper treatment of quantification in ordinary english. In *Approaches to Natural Language: Proc. of the 1970 Stanford Workshop on Grammar and Semantics*, 1973.
8. I. Pratt and A. Third. More fragments of language. *Notre Dame Journal of Formal Logic*, 2005.
9. L. Straßburger. What is a logic, and what is a proof? In J.-Y. Beziau, editor, *Logica Universalis*, pages 135–145. Birkhauser, 2005.
10. A. Third. *Logical Analysis of Fragments of Natural Language*. PhD thesis, Faculty of Engineering and Physical Sciences, University of Manchester, 2006.