Complexity Boundaries for Horn Description Logics

Markus Krötzsch and **Sebastian Rudolph** and **Pascal Hitzler**
Institute AIFB, Universität Karlsruhe, Germany
{mak,sru,phi}@aifb.uni-karlsruhe.de

Abstract

Horn description logics (Horn-DLs) have recently started to attract attention due to the fact that their (worst-case) data complexities are in general lower than their overall (i.e. combined) complexities, which makes them attractive for reasoning with large ABoxes. However, the natural question whether Horn-DLs also provide advantages for TBox reasoning has hardly been addressed so far. In this paper, we therefore provide a thorough and comprehensive analysis of the combined complexities of Horn-DLs. While the combined complexity for many Horn-DLs turns out to be the same as for their non-Horn counterparts, we identify subboolean DLs where Hornness simplifies reasoning.

Introduction

One of the driving motivations behind description logic (DL) research is to design languages which maximise the availability of expressive language features for the knowledge modelling process, while at the same time striving for the most inexpensive languages in terms of computational complexity. A particularly prominent case in point is the DL-based Web Ontology Language OWL,[1] which is a W3C recommended standard since 2004. OWL (more precisely, OWL DL) is indeed among the most expressive known knowledge representation languages which are also decidable.

Of particular interest for practical investigations are obviously tractable DLs. While not being boolean closed, and thus relatively inexpressive, they recently receive increasing attention as they promise to provide a good trade-off between expressivity and scalability (see e.g. (Baader, Brandt, & Lutz 2005)).

At the same time, Horn-DLs have been introduced (Grosof *et al.* 2003; Hustadt, Motik, & Sattler 2005), as their generally lower data complexities make them a natural and efficient choice for reasoning with large numbers of individuals, i.e. for ABox-reasoning. However, the natural question whether Horn-DLs also provide advantages for TBox reasoning – in terms of combined complexity – has hardly been addressed so far.

[1] http://www.w3.org/2004/OWL/

In this paper, we therefore provide a thorough and comprehensive analysis of the combined complexities of Horn-DLs. While the combined complexity for many Horn-DLs turns out to be the same as for their non-Horn counterparts – which is no surprise –, we are also able to identify subboolean DLs where the Hornness restriction improves reasoning complexity.

The paper is structured as follows. After recalling some preliminaries on DLs, we deal in turn with the Horn versions of $\mathcal{FL}_0$, $\mathcal{FL}^-$ and $\mathcal{FLE}$ and some of their variants. We will see that these provide us with a fairly complete picture of the complexities of Horn-DLs.

Full proofs have been omitted due to lack of space, but we made an effort to include proof sketches wherever possible. Full proofs can be found in (Krötzsch, Rudolph, & Hitzler 2007).

Preliminaries

In this section, we briefly recall some basic definitions of DLs and introduce our notation. We start with the rather expressive description logic $\mathcal{SHOIQ}\circ$ in order to define other DLs as restrictions thereof.

A $\mathcal{SHOIQ}\circ$ knowledge base (KB) is based on sets N_R (*role names*) N_C (*concept names*), and N_I (*individual names*). We define the set of $\mathcal{SHOIQ}\circ$ *atomic concepts* $\mathbf{C} := \mathsf{N}_C \cup \{\{i\} \mid i \in \mathsf{N}_I\}$. The set of $\mathcal{SHOIQ}\circ$ *(abstract) roles* is $\mathbf{R} = \mathsf{N}_R \cup \{R^- \mid R \in \mathsf{N}_R\}$. In the following, we leave this vocabulary implicit and assume that A, B, are atomic concepts, a, b, i are individual names, and R, S are abstract roles. Those can be used to define concept descriptions employing the constructors from the upper part of Table 1. We use C, D to denote concept descriptions.

A $\mathcal{SHOIQ}\circ$ knowledge base consists of three finite sets of axioms that are referred to as *RBox*, *TBox*, and *ABox*. The possible axiom types are displayed in the lower part of Table 1. [2]

[2] Setting $\mathsf{Inv}(R) = R^-$ and $\mathsf{Inv}(R^-) = R$, we put the common syntactical constraints on the RBox: it may contain axioms of the form $S \sqsubseteq R$ iff it also contains $\mathsf{Inv}(S) \sqsubseteq \mathsf{Inv}(R)$, and axioms of the form $\mathsf{Trans}(R)$ iff it also contains $\mathsf{Trans}(\mathsf{Inv}(R))$. By $\sqsubseteq^*$ we denote the reflexive-transitive closure of $\sqsubseteq$. A role R is *transitive* whenever there is a role S such that $\mathsf{Trans}(S)$, $R \sqsubseteq^* S$ and $S \sqsubseteq^* R$. R is *simple* if it has no transitive subroles, i.e., if $S \sqsubseteq^* R$ implies that S is not transitive. In the presence of role composition axioms, additional

Syntax	Semantics	
R^-	$\{(x,y) \mid (y,x) \in R^\mathcal{I}\}$	
$\top$	$\mathcal{D}$	
$\bot$	$\emptyset$	
$\{i\}$	$\{i^\mathcal{I}\}$	
$\neg C$	$\mathcal{D} \setminus C^\mathcal{I}$	
$C \sqcap D$	$C^\mathcal{I} \cap D^\mathcal{I}$	
$C \sqcup D$	$C^\mathcal{I} \cup D^\mathcal{I}$	
$\forall R.C$	$\{x \mid (x,y) \in R^\mathcal{I}$ implies $y \in C^\mathcal{I}\}$	
$\exists R.C$	$\{x \mid$ for some $y \in \mathcal{D},\ (x,y) \in R^\mathcal{I}, y \in C^\mathcal{I}\}$	
$\leq n\,R.C$	$\{x \mid \#\{y \in \mathcal{D} \mid (x,y) \in R^\mathcal{I}, y \in C^\mathcal{I}\} \leq n\}$	
$\geq n\,R.C$	$\{x \mid \#\{y \in \mathcal{D} \mid (x,y) \in R^\mathcal{I}, y \in C^\mathcal{I}\} \geq n\}$	
$S \sqsubseteq R$	$S^\mathcal{I} \subseteq R^\mathcal{I}$	RBox
$S_1 \circ \ldots \circ S_n \sqsubseteq R$	$S_1^\mathcal{I} \circ \ldots \circ S_n^\mathcal{I} \subseteq R^\mathcal{I}$	RBox
$\mathsf{Trans}(S)$	$S^\mathcal{I}$ is transitive	RBox
$C \sqsubseteq D$	$C^\mathcal{I} \subseteq D^\mathcal{I}$	TBox
$A(a)$	$a^\mathcal{I} \in A^\mathcal{I}$	ABox
$R(a,b)$	$(a^\mathcal{I}, b^\mathcal{I}) \in R^\mathcal{I}$	ABox
$a \approx b$	$a^\mathcal{I} = b^\mathcal{I}$	ABox

Table 1: Role/concept constructors and axiom types in $\mathcal{SHOIQ}\circ$. Semantics refers to an interpretation $\mathcal{I}$ with domain $\mathcal{D}$.

Mark that our ABoxes are *extensionally reduced*[3]. It is known that this does not restrict the expressivity of the logic since complex ABox statements can easily be moved into the TBox by introducing auxiliary concept names. Moreover, we do not explicitly consider concept/role equivalence $\equiv$, since it can be modelled via mutual concept/role inclusions.

We adhere to the common model-theoretic semantics for $\mathcal{SHOIQ}\circ$ with general concept inclusion axioms (GCIs): an interpretation $\mathcal{I}$ consists of a set $\mathcal{D}$ called *domain* together with a function $\cdot^\mathcal{I}$ mapping individual names to elements of $\mathcal{D}$, class names to subsets of $\mathcal{D}$, and role names to subsets of $\mathcal{D}\times\mathcal{D}$. This function is inductively extended to abstract roles and concept descriptions and finally used to decide whether the interpretation satisfies given axioms (according to Table 1).

Now we define the class of Horn DLs. This is done by first defining Horn-$\mathcal{SHOIQ}\circ$. For any sublogic $\mathcal{DL}$ of $\mathcal{SHOIQ}\circ$ we define Horn-$\mathcal{DL}$:= Horn-$\mathcal{SHOIQ}\circ \cap \mathcal{DL}$

Definition 1 The description logic Horn-$\mathcal{SHOIQ}\circ$ is defined as $\mathcal{SHOIQ}\circ$ except that the only allowed concept inclusions are of the form $\mathbf{C}_0^- \sqsubseteq \mathbf{C}_1^+$ according to the grammar in Table 2.

One can show that any Horn-$\mathcal{SHOIQ}\circ$ knowledge base can be transformed into an equisatisfiable Horn-$\mathcal{SHOIQ}\circ$ knowledge base containing GCIs only of the types $\top \sqsubseteq A$, $A \sqcap A' \sqsubseteq B$, $\exists R.A \sqsubseteq B$, $A \sqsubseteq \bot$, $A \sqsubseteq \exists R.B$, $A \sqsubseteq \forall S.B$, $A \sqsubseteq \geq n\,R.B$, and $A \sqsubseteq \leq 1\,R.B$, with A, A', B all being concept names. The respective proof involves the conversion

$$\mathbf{C}_1^+ \quad ::= \quad \top \mid \bot \mid \neg\mathbf{C}_1^- \mid \mathbf{C}_1^+ \sqcap \mathbf{C}_1^+ \mid \mathbf{C}_0^+ \sqcup \mathbf{C}_1^+ \mid \exists\mathbf{R}.\mathbf{C}_1^+ \mid \forall\mathbf{S}.\mathbf{C}_1^+ \mid \forall\mathbf{R}.\mathbf{C}_0^+ \mid \geq n\,\mathbf{R}.\mathbf{C}_1^+ \mid \leq 1\,\mathbf{R}.\mathbf{C}_0^- \mid \mathbf{A}$$

$$\mathbf{C}_1^- \quad ::= \quad \top \mid \bot \mid \neg\mathbf{C}_1^+ \mid \mathbf{C}_0^- \sqcap \mathbf{C}_1^- \mid \mathbf{C}_1^- \sqcup \mathbf{C}_1^- \mid \exists\mathbf{S}.\mathbf{C}_1^- \mid \exists\mathbf{R}.\mathbf{C}_0^- \mid \forall\mathbf{R}.\mathbf{C}_1^- \mid \geq 2\,\mathbf{R}.\mathbf{C}_0^- \mid \leq n\,\mathbf{R}.\mathbf{C}_1^+ \mid \mathbf{A}$$

$$\mathbf{C}_0^+ \quad ::= \quad \top \mid \bot \mid \neg\mathbf{C}_0^- \mid \mathbf{C}_0^+ \sqcap \mathbf{C}_0^+ \mid \mathbf{C}_0^+ \sqcup \mathbf{C}_0^+ \mid \forall\mathbf{R}.\mathbf{C}_0^+$$

$$\mathbf{C}_0^- \quad ::= \quad \top \mid \bot \mid \neg\mathbf{C}_0^+ \mid \mathbf{C}_0^- \sqcap \mathbf{C}_0^- \mid \mathbf{C}_0^- \sqcup \mathbf{C}_0^- \mid \exists\mathbf{R}.\mathbf{C}_0^- \mid \mathbf{A}$$

Table 2: A grammar for defining Horn-$\mathcal{SHOIQ}\circ$. **A**, **R**, and **S** denote the sets of all atomic concepts, abstract roles, and simple roles, respectively. The presentation is slightly simplified by exploiting associativity and commutativity of $\sqcap$ and $\sqcup$, and by omitting $\geq 1\,R.C$ if $\exists R.C$ is present.

to negation normal form, the elimination of disjunction and negation by exploiting the Horn structure, and the introduction of new concept names to substitute subexpressions. The main argumentation is similar to *Lloyd-Topor transformations* that are considered in logic programming.

Finally, we observe that the following standard reasoning are mutually reducuble even when restricting to Horn knowledge bases.

Knowledge base satisfiability. We call a knowledge base *satisfiable*, if it has a model, i.e., if there exists an interpretation $\mathcal{I}$ satisfying all axioms of the knowledge base.

Instance checking. For a given individual a and a given concept description C of form $\mathbf{C}_0^-$, we ask whether $C(a)$ is satisfied in all models of the knowledge base KB. This task can be reduced to the knowledge base satisfiability problem in the following way: Letting A be a new, unused concept name, check whether the knowledge base KB$\cup\{A(a), A \sqcap C \sqsubseteq \bot\}$ is unsatisfiable.

Entailment of TBox axioms. A TBox axiom (GCI) $C \sqsubseteq D$ is *entailed* by a knowledge base KB if it is satisfied by all interpretations that satisfy the knowledge base. If C is of the form $\mathbf{C}_1^+$ and D is of the form $\mathbf{C}_0^-$, this problem can be reduced to the instance checking problem: let A, B be concept names not already present in the knowledge base KB and a be a new individual name. Then instance check for $B(a)$ in KB$\cup\{A \sqsubseteq C, D \sqsubseteq B, A(a)\}$.

Concept satisfiability. A concept description C is *satisfiable* (with respect to a given knowledge base) if the knowledge base has a model $\mathcal{I}$ with $C^\mathcal{I} \neq \emptyset$. If C has the form $\mathbf{C}_1^+$, this can be reduced to the preceding problem by checking whether $C \sqsubseteq \bot$ is entailed by the considered knowledge base.

Hence, we have shown that all reasoning problems can be reduced to knowledge base satisfiability. Querying a knowledge base for some statement is equivalent to checking whether the negation of this statement entails unsatisfiability, which explains why the above (Horn) restrictions on queries are in a sense dual to the restrictions on Horn axioms.

Horn-$\mathcal{FL}_0$

The description logic $\mathcal{FL}_0$ is indeed very simple: $\top$, $\bot$, $\sqcap$, and $\forall$ are the only operators allowed. Yet, checking the

restrictions apply to ensure decidability and the definition of simple roles has to be modified. See (Horrocks, Kutz, & Sattler 2006) for a thorough treatise.

[3] I.e., only atomic concepts occur in the ABox.

$A \sqsubseteq C$	$\top \sqsubseteq C$	$A(c)$	$R \sqsubseteq T$
$A \sqcap B \sqsubseteq C$	$A \sqsubseteq \bot$	$R(c, d)$	$R \circ S \sqsubseteq T$
$A \sqsubseteq \forall R.C$		$c \approx d$	

Table 3: Normal form for Horn-$\mathcal{FL}_0{}^+$. A, B, and C are names of atomic concepts or nominal classes, R, S, and T (possibly inverse) role names, and c and d individual names.

satisfiability of $\mathcal{FL}_0$ knowledge bases is already ExpTime-complete (Baader, Brandt, & Lutz 2005). In this section, we show that Horn-$\mathcal{FL}_0$ is in P, and thus is much simpler than its non-Horn counterpart. In fact, we can even extend this logic with additional expressive means without sacrificing tractability.

Definition 2 The description logic $\mathcal{FL}_0{}^+$ is the extension of $\mathcal{FL}_0$ with nominals, role hierarchies, role composition, and inverse roles. The Horn fragment of $\mathcal{FL}_0{}^+$ is denoted Horn-$\mathcal{FL}_0{}^+$.

To show that Horn-$\mathcal{FL}_0{}^+$ is in P, we will reduce satisfiability checking for Horn-$\mathcal{FL}_0{}^+$ to satisfiability checking in the 3-variable fragment of function-free Horn logic. A Horn-clause is a disjunction of atomic formulae or negations thereof, which contains at most one non-negated atom, and with all variables quantified universally. Horn-clauses are commonly written as implications (with possibly empty head or body), and without explicitly specifying the quantifiers. The following is straightforward.

Proposition 3 Satisfiability of a logical theory that consists of function-free Horn-clauses with a bounded number of variables can be checked in time polynomial w.r.t. the size of the theory.

We say that a Horn-$\mathcal{FL}_0{}^+$ knowledge base is in *normal form* if it contains only axioms of the forms shown in Table 3. The simple algorithm for establishing the following is detailed in (Krötzsch, Rudolph, & Hitzler 2007).

Lemma 4 Checking satisfiability of a Horn-$\mathcal{FL}_0{}^+$ knowledge base can be reduced in linear time to checking satisfiability of a Horn-$\mathcal{FL}_0{}^+$ knowledge base that is in normal form.

The normal form transformation is necessary to ensure that at most three distinct variables are needed within the first-order version of every axiom. Since the equality predicate can be axiomatised in function-free Horn-logic, we find that every Horn-$\mathcal{FL}_0{}^+$ knowledge base in normal form is semantically equivalent to a logical theory in the 3-variable fragment of function-free Horn-logic. Summing up, we obtain the following.

Theorem 5 Satisfiability for Horn-$\mathcal{FL}_0{}^+$ knowledge bases can be decided in polynomial time.

It is not hard to see that the well-known DLP-fragment of $\mathcal{SHIQ}$ (Grosof *et al.* 2003) does indeed allow for a similar reduction to 3-variable Horn logic, and thus has an at most polynomial time complexity. To the best of our knowledge, this result has not been spelled out before.

Horn-$\mathcal{FL}^-$

Extending Horn-$\mathcal{FL}_0$ quickly leads to intractable logics. As we will see below, all logics between Horn-$\mathcal{FL}^-$ and Horn-$\mathcal{FLOH}^-$ are PSpace-complete. As shown in Theorem 11, adding more expressivity to Horn-$\mathcal{FL}^-$ increases the complexity even further.

Definition 6 The description logic $\mathcal{FL}^-$ is the extension of $\mathcal{FL}_0$ with concept expressions of the form $\exists R.\top$, and $\mathcal{FLOH}^-$ is the extension of $\mathcal{FL}^-$ with nominals and role hierarchies. The respective Horn fragments are denoted Horn-$\mathcal{FL}^-$ and Horn-$\mathcal{FLOH}^-$.

To show the claimed result, one first has to show that Horn-$\mathcal{FL}^-$ is PSpace-hard. This is done in a straightforward way by reducing the halting problem of deterministic Turing machines with polynomially bounded storage. We omit this proof here, and sketch a more interesting Turing machine reduction in the next section instead. Interested readers will find all details spelled out in (Krötzsch, Rudolph, & Hitzler 2007).

Showing that Horn-$\mathcal{FLOH}^-$ is contained in PSpace turns out to be more involved, and we take some time to provide an extended sketch of the full proof. A typical approach would be to describe a tableau algorithm, and to show that a nondeterministic "depth-first" search can detect a clash using only polynomial memory. This is complicated in Horn-$\mathcal{FLOH}^-$ in two ways. Firstly, the length of a computation path can easily be exponential, so that not even a single path can be stored in PSpace. Secondly, the presence of nominals changes the structure of the tableau by enabling loops, which occur whenever an element is inferred to belong to a nominal concept. To deal with those problems, the algorithm needs to eagerly forget its previous computation path, and use massive nondeterminism for still finding the correct derivations. It turns out that this can best be achieved by computing *backwards*.

The following algorithm assumes the knowledge base to be in a normal form, similar to the one introduced in Table 3. We omit the details for reasons of space.

Definition 7 Given a Horn-$\mathcal{FLOH}^-$ knowledge base KB, a set of relevant concept expressions is defined as $\mathsf{cl}(KB) = \mathbf{C} \cup \{QR.C | R \in \mathbf{R}, C \in \mathbf{C}, Q \in \{\exists, \forall\}\} \cup \{\top, \bot\}$. Let I denote the set $\mathbf{I} \cup \{a, b\}$, where a, b are fresh names, and define a set $\mathcal{T}_I$ of expressions as $\mathcal{T}_I := \{C(e) \mid C \in \mathsf{cl}(KB), e \in I\} \cup \{R(e, f) \mid R \in \mathbf{R},\ e, f \in I\}$.

The algorithm nondeterministically selects one element $g \in I$, and initialises $T \subseteq \mathcal{T}_I$ by setting $T := \{\bot(g)\}$. Let $T_{e \mapsto f}$ abbreviate the set $\{C(f) \mid C(e) \in T\} \cup \{R(f, g) \mid R(e, g) \in T,\ g \in I\} \cup \{R(g, f) \mid R(g, e) \in T,\ g \in I\}$ and define $T_e := T_{e \mapsto e}$. The algorithm repeatedly modifies T by nondeterministically applying one of the following rules:

(N1) Given any $X \in \mathcal{T}_I$, set $T := T \cup \{X\}$.

1. $T := T \cup \{\top(e)\}$
2. if $e \in \mathbf{I}$ is a named individual, $T := T \cup \{\{e\}(e)\}$
3. for each $A(e) \in KB$, $T := T \cup \{A(e)\}$
4. for each $R(e, f) \in KB$, $T := T \cup \{R(e, f)\}$
5. for each $e \approx f \in KB$, $T := T \cup \{\{f\}(e), \{e\}(f)\}$
6. for each $\{f\}(e) \in T$, do the following
   (a) for each $C(f) \in T$, $T := T \cup \{C(e)\}$
   (b) for each $g \in I$ and $R(f, g) \in T$, $T := T \cup \{R(e, g)\}$
   (c) for each $g \in I$ and $R(g, f) \in T$, $T := T \cup \{R(g, e)\}$
   (d) for each $C(e) \in T$, $T := T \cup \{C(f)\}$
   (e) for each $g \in I$ and $R(e, g) \in T$, $T := T \cup \{R(f, g)\}$
   (f) for each $g \in I$ and $R(g, e) \in T$, $T := T \cup \{R(g, f)\}$
7. for each $A \sqsubseteq C \in KB$, if $A(e) \in T$ then $T := T \cup \{C(e)\}$
8. for each $A \sqcap B \sqsubseteq C \in KB$, if $A(e) \in T$ and $B(e) \in T$ then $T := T \cup \{C(e)\}$
9. for each $R \sqsubseteq S \in KB$, do the following:
   (a) for each $f \in I$, if $R(e, f) \in T$, then $T := T \cup \{S(e, f)\}$
   (b) if $\exists R.\top(e) \in T$ then $T := T \cup \{\exists S.\top(e)\}$
10. for each $f \in I$ and $R(e, f) \in T$ with $R(e, f)$ not inactive, $T := T \cup \{\exists R.\top(e)\}$
11. for each $\forall R.C(e) \in T$ and $R(e, g) \in T, T := T \cup \{C(g)\}$

Table 4: Derivation rules for Horn-$\mathcal{FLOH}^-$.

(N2) If there is some individual $e \in I$ and $X \in T$ such that X can be derived from $T \setminus \{X\}$ using one of the rules in Table 4, set $T := T \setminus \{X\}$.

(N3) If $T_a = \{R(e, a)\}$ for some $e \in I \setminus \{a\}$ such that $\exists R.\top(e) \in T$, set $T := T \setminus T_a$.

(N4) If $T_a = \emptyset$, set $T := (T \cup T_{b \mapsto a}) \setminus T_b$.

(N5) If $T = \emptyset$, return "unsatisfiable."

It can be shown that the above algorithm can indeed detect unsatisfiability of Horn-$\mathcal{FLOH}^-$ knowledge bases in polynomial space. Intuitively, the algorithm guesses an assumed clash $\bot(e)$ within a "minimal" tableau, and tries to reconstruct a derivation procedure for this clash. The current set of unproved assumptions is given by T, and additional assumptions might be made in step (N1). The assumptions might refer to a known individual, or to a or b, which are placeholders for unnamed elements that occur in a proper forward tableau construction. In (N2), existing assumptions that can be shown by deduction are deleted. Rule (N3) is similar, but accounts for the special case that the existence of a new, anonymous individual was inferred. (N4) merely allows to rename b into a if a is currently unused, while (N5) checks if all assumptions have been reduced successfully.

In spite of the possible size of the tableau that is explored by the algorithm, it suffices to consider at most two anonymous individuals (a and b) in each step. The strategy for guiding nondeterminism is to eagerly reduce assumptions on a, and to introduce (at most one) role predecessor b of a if needed. When all assumptions on a have been reduced, b is copied to a and reduction continues. If a and b are empty, assumptions on named individuals are considered. The success of the algorithm hinges upon the fact that it nondeterministically guesses from which premises each statement can be inferred, and that it reduces assertions in an appropriate order.

Since NPSpace is well-known to coincide with PSpace, we obtain the following result.

Theorem 8 Deciding knowledge base satisfiability in any description logic between Horn-$\mathcal{FL}^-$ and Horn-$\mathcal{FLOH}^-$ is PSpace-complete.

Horn-$\mathcal{FLE}$

$\mathcal{FLE}$ further extends $\mathcal{FL}^-$ by allowing arbitrary existential role quantifications, which turns out to raise the complexity of Horn-$\mathcal{FLE}$ to ExpTime. Note that inclusion in ExpTime is obvious since $\mathcal{FLE}$ is a fragment of $\mathcal{SHIQ}$ which is also in ExpTime (Tobies 2001). To show that Horn-$\mathcal{FLE}$ is ExpTime-hard, we reduce the halting problem of polynomially space-bounded alternating Turing machines, defined next, to the concept subsumption problem. An extended discussion of the core proof is found in (Krötzsch, Rudolph, & Hitzler 2007).

Definition 9 An *alternating Turing machine* (ATM) $\mathcal{M}$ is a tuple (Q, Σ, Δ, q_0) where

- $Q = U \dot{\cup} E$ is the disjoint union of a finite set of *universal states* U and a finite set of *existential states* E,
- Σ is a finite *alphabet* that includes a *blank* symbol $\square$,
- $\Delta \subseteq (Q \times \Sigma) \times (Q \times \Sigma \times \{l, r\})$ is a *transition relation*, and
- $q_0 \in Q$ is the *initial state*.

A (universal/existential) *configuration* of $\mathcal{M}$ is a word $\alpha \in \Sigma^* Q \Sigma^*$ ($\Sigma^* U \Sigma^* / \Sigma^* E \Sigma^*$). A configuration α' is a *successor* of a configuration α if one of the following holds:

1. $\alpha = w_l q \sigma \sigma_r w_r$, $\alpha' = w_l \sigma' q' \sigma_r w_r$, and $(q, \sigma, q', \sigma', r) \in \Delta$,
2. $\alpha = w_l q \sigma$, $\alpha' = w_l \sigma' q' \square$, and $(q, \sigma, q', \sigma', r) \in \Delta$,
3. $\alpha = w_l \sigma_l q \sigma w_r$, $\alpha' = w_l q' \sigma_l \sigma' w_r$, and $(q, \sigma, q', \sigma', l) \in \Delta$,

where $q \in Q$ and $\sigma, \sigma', \sigma_l, \sigma_r \in \Sigma$ as well as $w_l, w_r \in \Sigma^*$. Given some natural number s, the possible *transitions in space s* are defined by additionally requiring that $|\alpha'| \leq s+1$.

The set of *accepting configurations* is the least set which satisfies the following conditions. A configuration α is accepting iff

- α is a universal configuration and all its successor configurations are accepting, or
- α is an existential configuration and at least one of its successor configurations is accepting.

Note that universal configurations without any successors here play the rôle of accepting final configurations, and thus form the basis for the recursive definition above. $\mathcal{M}$ *accepts* a given word $w \in \Sigma^*$ (in space s) iff the configuration $q_0 w$ is accepting (when restricting to transitions in space s).

This definition is inspired by the complexity classes NP and co-NP, which are characterised by non-deterministic Turing machines that accept an input if either at least one or all possible runs lead to an accepting state. An ATM can switch between these two modes and indeed turns out to be more powerful than classical Turing machines of either kind. In particular, ATMs can solve ExpTime problems

(1) Left and right transition rules:

$A_q \sqcap H_i \sqcap C_{\sigma,i} \sqsubseteq \exists S_\delta.(A_{q'} \sqcap H_{i+1} \sqcap C_{\sigma',i})$ with $\delta = (q, \sigma, q', \sigma', r)$, $i < p(|w|) - 1$

$A_q \sqcap H_i \sqcap C_{\sigma,i} \sqsubseteq \exists S_\delta.(A_{q'} \sqcap H_{i-1} \sqcap C_{\sigma',i})$ with $\delta = (q, \sigma, q', \sigma', l)$, $i > 0$

(2) Memory: $H_j \sqcap C_{\sigma,i} \sqsubseteq \forall S_\delta.C_{\sigma,i} \quad i \neq j$

(3) Existential acceptance: $A_q \sqcap \exists S_\delta.A \sqsubseteq A$ for all $q \in E$

(4) Universal acceptance:

$A_q \sqcap H_i \sqcap C_{\sigma,i} \sqcap \bigsqcap_{\delta \in \tilde{\Delta}} (\exists S_\delta.A) \sqsubseteq A$ with $q \in U$,
$\tilde{\Delta} = \{(q, \sigma, q', \sigma', x) \in \Delta\}$, $x \in \{r \mid i < p(|w|) - 1\} \cup \{l \mid i > 0\}$

Table 5: Knowledge base $K_{\mathcal{M},w}$ simulating a polynomially space-bounded ATM. The rules are instantiated for all $q, q' \in Q$, $\sigma, \sigma' \in \Sigma$, $i, j \in \{0, \ldots, p(|w|) - 1\}$, and $\delta \in \Delta$.

in polynomial space: the complexity class APSpace of languages accepted by polynomially space-bounded ATMs coincides with the complexity class ExpTime (Chandra, Kozen, & Stockmeyer 1981). As usual, a language L is accepted by a polynomially space-bounded ATM iff there is a polynomial p such that, for every word $w \in \Sigma^*$, $w \in L$ iff w is accepted in space $p(|w|)$. In the following, we exclusively deal with polynomially space-bounded ATMs, and so we omit additions such as "in space s" when clear from the context.

In the following, we consider a fixed ATM $\mathcal{M}$ denoted as in Definition 9, and a polynomial p that defines a bound for the required space. For any word $w \in \Sigma^*$, we construct a Horn-$\mathcal{FLE}$ knowledge base $K_{\mathcal{M},w}$ and show that acceptance of w by the ATM $\mathcal{M}$ can be reduced to checking concept subsumption. Intuitively, the elements of an interpretation domain of $K_{\mathcal{M},w}$ represent possible configurations of $\mathcal{M}$, encoded by the following concept names:

- A_q for $q \in Q$: the ATM is in state q,
- H_i for $i = 0, \ldots, p(|w|) - 1$: the ATM is at position i on the storage tape,
- $C_{\sigma,i}$ with $\sigma \in \Sigma$ and $i = 0, \ldots, p(|w|) - 1$: position i on the storage tape contains symbol σ,
- A: the ATM accepts this configuration.

This approach is pretty standard, and it is not too hard to axiomatise a successor relation S and appropriate acceptance conditions in $\mathcal{ALC}$. But, as explained in (Krötzsch, Rudolph, & Hitzler 2006), this reduction is not applicable in Horn-$\mathcal{FLE}$. Hence, we reduce the halting problem to concept subsumption, and adapt the recursively defined acceptance condition of Definition 9 to ensure that the initial state must be accepting in *all* possible models. Also, we encode individual transitions by using a distinguished successor relation for each translation in Δ.

Now consider the knowledge base $K_{\mathcal{M},w}$ given in Table 5. The roles S_δ, $\delta \in \Delta$, describe a configuration's successors using the translation δ. The initial configuration for word w is described by the concept expression I_w which is set to

$$A_{q_0} \sqcap H_0 \sqcap C_{\sigma_0,0} \sqcap \ldots \sqcap C_{\sigma_{|w|-1},|w|-1} \sqcap C_{\square,|w|} \sqcap \ldots \sqcap C_{\square,p(|w|)-1}$$

where σ_i denotes the symbol at the ith position of w. It is easy to see that $K_{\mathcal{M},w}$ and $I_w \sqsubseteq A$ are in Horn-$\mathcal{FLE}$. Checking whether the initial configuration is accepting is equivalent to checking whether $I_w \sqsubseteq A$ follows from $K_{\mathcal{M},w}$. Based on the above knowledge base, the following result was established in (Krötzsch, Rudolph, & Hitzler 2006).

Theorem 10 Checking concept subsumption in any description logic between Horn-$\mathcal{FLE}$ and Horn-$\mathcal{SHIQ}$ is ExpTime-complete.

Note that, even in Horn logics, it is straightforward to reduce knowledge base satisfiability to the entailment of the concept subsumption $\top \sqsubseteq \bot$. The proof that was used to establish the previous result is suitable for obtaining further complexity results for logical fragments that are not above Horn-$\mathcal{FLE}$.

Theorem 11 (a) Let $\mathcal{EL}^{\leq 1}$ denote $\mathcal{EL}$ extended with number restrictions of the form $\leq 1\, R.\top$.

(b) Let $\mathcal{FL}\circ^-$ denote $\mathcal{FL}^-$ extended with composition of roles.

(c) Let $\mathcal{FLI}^-$ denote $\mathcal{FL}^-$ extended with inverse roles.

Horn-$\mathcal{FL}\circ^-$ is ExpTime-hard, and both Horn-$\mathcal{EL}^{\leq 1}$ and Horn-$\mathcal{FLI}^-$ are ExpTime-complete.

Proof. The results are established by modifying the knowledge base $K_{\mathcal{M},w}$ to suite the given fragment. We restrict to providing the required modifications; the full proofs are analogous to the proof for Horn-$\mathcal{FLE}$.

(a) Replace axioms (2) in Table 5 with the following statements:

$$\top \sqsubseteq \leq 1\, S_\delta.\top \qquad H_j \sqcap C_{\sigma,i} \sqcap \exists S_\delta.\top \sqsubseteq \exists S_\delta.C_{\sigma,i},\ i \neq j$$

(b) Replace axioms (1) with axioms of the form

$$A_q \sqcap H_i \sqcap C_{\sigma,i} \sqsubseteq \exists S_\delta.\top \sqcap \forall S_\delta.(A_{q'} \sqcap H_{i\pm 1} \sqcap C_{\sigma',i}).$$

Any occurrence of concept A is replaced by $\exists R_A.\top$, with R_A a new role. Moreover, we introduce roles $R_{A\delta}$ for each transition δ, and replace any occurrence of $\exists S_\delta.A$ with $\exists R_{A\delta}.\top$. Finally, the following axioms are added:

$$S_\delta \circ R_A \sqsubseteq R_{A\delta} \quad \text{for each } \delta \in \Delta.$$

(c) Axioms (1) are replaced as in (b). Any occurrence of $\exists S_\delta.A$ is now replaced with a new concept name $A_{S\delta}$, and the following axioms are added:

$$A \sqsubseteq \forall S_\delta^{-1}.A_{S\delta} \quad \text{for each } \delta \in \Delta.$$

It is easy to see that those changes still enable analogous reductions. Inclusion results for Horn-$\mathcal{EL}^{\leq 1}$ and Horn-$\mathcal{FLI}^-$ are immediate from their inclusion in $\mathcal{SHIQ}$. □

Summary

Horn logics, while having a long tradition in logic programming, have only recently been studied in the context of description logics, mainly due to their lower data complexities (Hustadt, Motik, & Sattler 2005). In this work, we have investigated the effects of Hornness on the overall complexity of DL reasoning, and we have shown that only the Horn

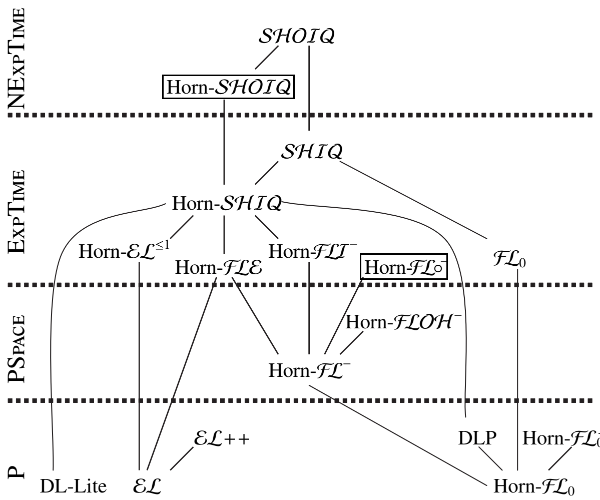

Figure 1: Overview over complexity boundaries for combined complexity of (Horn) description logics in the presence of GCIs. The exact location of the boxed DLs is conjectured, and has not been shown yet.

fragments of certain subboolean description logics are actually less complex than their non-Horn versions. On the other hand, the well-known tractable DLs $\mathcal{EL}$++ (Baader, Brandt, & Lutz 2005), DLP (Grosof *et al.* 2003), and DL-Lite (Calvanese *et al.* 2005) are also recognised as (fragments of) Horn-logics. We thus obtain a unified picture of combined complexities of some of the most important tractable DLs currently discussed.[4] The main results of our work are summarised in Figure 1.

While most of the displayed relationships have been verified above, Figure 1 also includes two open conjectures that are left to future research: Horn-$\mathcal{SHOIQ}$ might also turn out to be in ExpTime, whereas Horn-$\mathcal{FLo}^{-}$ could even be NExpTime-hard. In addition, some related results have not been included in Figure 1. In particular, we have shown that (Horn) disjunction and atomic negation never increase the complexity of Horn-logics. Accordingly, the extension of $\mathcal{EL}$++ to Horn-$\mathcal{ELU}^{(\neg)}$++ is still tractable, while it was shown in (Baader, Brandt, & Lutz 2005) that both $\mathcal{ELU}$ and $\mathcal{EL}^{(\neg)}$ are ExpTime-complete. An interesting application of this extension is the use of DLs as query languages, since the use of disjunctions is not constrained within queries (which are treated as negated axioms) at all. This is exploited, for instance, in the semantic search implementation of Semantic MediaWiki (Völkel *et al.* 2006), which indeed supports a (syntactically adopted) fragment of Horn-$\mathcal{ELU}$++ for querying large scale knowledge bases.

Our results on Horn-$\mathcal{FLE}$ and Horn-$\mathcal{EL}^{\leq 1}$ sharpen the known results on the non-extendibility of $\mathcal{EL}$. On the other hand, various expressive extensions that are known not to increase the complexity of $\mathcal{EL}$ were also shown to be tolerable in the case of Horn-$\mathcal{FL}_0$ and Horn-$\mathcal{FL}^{-}$. In particular, nominals and role hierarchies have had no negative effect on the worst-case complexity of any of the investigated Horn-logics. Yet, it is apparent from the technically more involved proof of Horn-$\mathcal{FLOH}^{-}$'s PSpace-completeness that especially nominals can easily make the reasoning task more complicated. As all other proofs, this proof is established directly, without referring to existing complexity results. While this is often increasing the length of the required argumentation, we believe that direct proofs are often most instructive for analysing the source of increased complexities.

[4] http://owl1_1.cs.manchester.ac.uk/tractable.html

Acknowledgements. We would like to thank the anonymous reviewers for helpful comments for improving this work. Research reported in this paper was supported by the EU in the IST projects X-Media (IST-FP6-026978) and NeOn (IST-2006-027595).

References

Baader, F.; Brandt, S.; and Lutz, C. 2005. Pushing the EL envelope. In *Proc. 19th Int. Joint Conf. on Artificial Intelligence (IJCAI-05)*. Edinburgh, UK: Morgan-Kaufmann Publishers.

Calvanese, D.; Giacomo, G. D.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2005. DL-Lite: Tractable description logics for ontologies. In Veloso, M. M., and Kambhampati, S., eds., *Proc. AAAI-05, Pittsburgh, Pennsylvania, USA*, 602–607. AAAI Press AAAI Press / The MIT Press.

Chandra, A. K.; Kozen, D. C.; and Stockmeyer, L. J. 1981. Alternation. *J. ACM* 28(1):114–133.

Grosof, B.; Horrocks, I.; Volz, R.; and Decker, S. 2003. Description logic programs: Combining logic programs with description logics. In *Proc. of WWW 2003, Budapest, Hungary, May 2003*, 48–57. ACM.

Horrocks, I.; Kutz, O.; and Sattler, U. 2006. The even more irresistible $\mathcal{SROIQ}$. In *Proc. of the 10th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR2006)*, 57–67. AAAI Press.

Hustadt, U.; Motik, B.; and Sattler, U. 2005. Data complexity of reasoning in very expressive description logics. In *Proc. 18th Int. Joint Conf. on Artificial Intelligence (IJCAI-05)*, 466–471. Edinburgh, UK: Morgan-Kaufmann Publishers.

Krötzsch, M.; Rudolph, S.; and Hitzler, P. 2006. On the complexity of Horn description logics. In *Proc. 2nd Workshop OWL Experiences and Directions*.

Krötzsch, M.; Rudolph, S.; and Hitzler, P. 2007. Complexity of Horn description logics. Technical report, Universität Karlsruhe (TH), Germany, Feb 2007.

Tobies, S. 2001. *Complexity Results and Practical Algorithms for Logics in Knowledge Representation*. Ph.D. Dissertation, RWTH Aachen, Germany.

Völkel, M.; Krötzsch, M.; Vrandečić, D.; Haller, H.; and Studer, R. 2006. Semantic Wikipedia. In *Proc. of the 15th International WWW Conference*.