A Van Benthem Theorem for Horn Description and Modal Logic

Fabio Papacchini and Frank Wolter

University of Liverpool, UK, {papacchf,wolter}@liverpool.ac.uk

Abstract. We introduce Horn simulations between interpretations and show that a first-order formula is equivalent to a Horn-$\mathcal{ALC}$ concept iff it is preserved under Horn simulations. Preservation under global Horn simulations characterizes first-order sentences equivalent to Horn-$\mathcal{ALC}$ TBoxes. We use Horn simulations to show that not every $\mathcal{ALC}$ concept (TBox) equivalent to a Horn FO formula is equivalent to a Horn-$\mathcal{ALC}$ concept (respectively TBox). For Horn modal logic, we thus solve an open problem posed by Sturm.

1 Introduction

Horn Description logics (Horn DLs) have been introduced as syntactically defined fragments of expressive description logics that fall within the Horn fragment of first-order logic and for which query evaluation is in PTime in data complexity [17, 18]. Since their introduction in 2005, there has been significant progress in developing reasoning algorithms for Horn DLs [19, 20], query evaluation using Horn DLs [12, 27], and in understanding the computational complexity of Horn DLs [21]. It also turned out that basic questions relevant for OBDA applications, such as query-inseparability and conservative extensions, FO-rewritability, query emptiness, ontology materialization, and query by example, admit more elegant solutions and are easier to solve computationally in Horn DLs than in the classical case [2, 5, 6, 13, 14]. The relationship between Horn DLs and PTime query evaluation is now also well understood [25, 16]. In contrast, the model theory of Horn DLs remains largely undeveloped.

In this paper, we make a first step towards a better understanding of the semantics of Horn DLs by providing a model-theoretic characterization of the expressive power of Horn-$\mathcal{ALC}$, the Horn fragment of the basic expressive DL $\mathcal{ALC}$. We introduce *Horn simulations* between interpretations (or, equivalently, *Horn simulation games*) and show that a first-order (FO) formula $\varphi(x)$ is equivalent to a Horn-$\mathcal{ALC}$ concept iff it is preserved under Horn simulations. We thus give a van Benthem style characterization of Horn-$\mathcal{ALC}$ concepts relative to FO. In fact, because of van Benthem's characterization of $\mathcal{ALC}$ concepts as the bisimulation invariant fragment of FO [4], it suffices to prove that an $\mathcal{ALC}$ concept is equivalent to a Horn-$\mathcal{ALC}$ concept iff it is preserved under Horn simulations. We also characterize Horn-$\mathcal{ALC}$ TBoxes: an $\mathcal{ALC}$ TBox is equivalent to a Horn-$\mathcal{ALC}$ TBox iff it is preserved under global Horn simulations. Again the more general

statement that a FO sentence is equivalent to a Horn-$\mathcal{ALC}$ TBox iff it is preserved under global Horn simulations follows from the characterization of $\mathcal{ALC}$ TBoxes as the fragment of FO that is invariant under global bisimulations and disjoint unions [23].

Horn simulations differ from standard bisimulations in at least two respects: they are non-symmetric (thus, we do not talk about invariance but preservation) and they relate sets to points (rather than points to points). They also employ as a 'subgame' the standard simulation game between interpretations known to characterize $\mathcal{ELU}$ ($\mathcal{EL}$ with 'union') concepts [23]. The definition of Horn simulations is inspired by games used to provide van Benthem style characterizations of concepts in weak DLs such as $\mathcal{FL}^-$ [22].

A modal variant of Horn-$\mathcal{ALC}$ was introduced earlier, and independently, in [29] with the explicit aim of giving a syntactic description of the intersection of Horn FO logic and modal logic. Here, we use Horn simulations to prove that Horn-$\mathcal{ALC}$ concepts (and, therefore, Horn modal logic) do *not* have this property: there are $\mathcal{ALC}$ concepts that are equivalent to Horn FO formulas but not equivalent to any Horn-$\mathcal{ALC}$ concept. A similar result is proved for Horn-$\mathcal{ALC}$ TBoxes using global Horn simulations.

Our proofs are inspired by Otto's finitary proofs of (extensions of) van Benthem's bisimulation characterization of modal logic via finitary bisimulations [28]: we introduce finitary Horn simulations characterizing the expressive power of Horn-$\mathcal{ALC}$ concepts (and TBoxes) of a fixed quantifier depth and show that these characterizations can be upgraded to Horn simulations. The proof also provides some insight into the relationship between preservation under products and preservation under Horn simulations. A proof via the original infinitary method using compactness and saturated models can be given as well, with one important caveat: it seems that this method only yields a characterization in which the sets selected in the Horn simulation game have to be FO-definable.

The structure of the paper is as follows. In Section 2, we introduce the relevant notation for description logic. We introduce simulations and Horn simulations and their finitary approximations. We also show that depth k Horn-$\mathcal{ALC}$ concepts are characterized by preservation under k-Horn simulations. In Section 3, we prove the chacterization of Horn-$\mathcal{ALC}$ concepts via Horn-simulations and show that not all $\mathcal{ALC}$ concepts equivalent to Horn FO formulas are equivalent to Horn-$\mathcal{ALC}$ concepts. In Section 4, we prove the characterization of Horn-$\mathcal{ALC}$ TBoxes via global Horn simulations and show that not all $\mathcal{ALC}$ TBoxes equivalent to Horn FO sentences are equivalent to Horn-$\mathcal{ALC}$ TBoxes. We close with a brief outlook.

2 Preliminaries

We use standard notation for description logics [1, 3]. For our characterization results, it is convenient to work with *finite* and disjoint sets $\mathsf{N_C}$ and $\mathsf{N_R}$ of *concept* and *role names*. Then $\mathcal{ALC}$ *concepts* C are defined by the following rule

$$C, C' := \quad A \mid \top \mid \bot \mid \neg C \mid C \sqcup C' \mid C \sqcap C' \mid C \to C' \mid \exists r.C \mid \forall r.C$$

where $A \in \mathsf{N_C}$ and $r \in \mathsf{N_R}$. An $\mathcal{ALC}$ *concept inclusion (CI)* takes the form $C \sqsubseteq D$, where C, D are $\mathcal{ALC}$ concepts. An $\mathcal{ALC}$ *TBox* is a finite set of $\mathcal{ALC}$ CIs. The *depth* of a concept C is the maximal number of nestings of $\exists r$ and $\forall r$ in C. For example, the concepts $\exists r.\exists r.A$ and $\exists r.\forall r.A$ have depth two. The *depth* of a TBox is the maximum over the depths of the concepts that occur in it. $\mathcal{ALC}$ concepts constructed from concept names using the constructors $\top$, $\sqcap$, $\sqcup$, and $\exists r.C$ only are called $\mathcal{ELU}$ *concepts*. As usual, $\mathcal{ELU}$ concepts not using $\sqcup$ are called $\mathcal{EL}$ *concepts*. Then *Horn-$\mathcal{ALC}$ concepts* R are defined by the following rule

$$R, R' ::= \bot \mid \top \mid \neg A \mid A \mid R \sqcap R' \mid L \to R \mid \exists r.R \mid \forall r.R$$

where $A \in \mathsf{N_C}$, $r \in \mathsf{N_R}$, and L is an $\mathcal{ELU}$ concept. A *Horn-$\mathcal{ALC}$ CI* takes the form $\top \sqsubseteq R$ with R a Horn-$\mathcal{ALC}$ concept. A *Horn-$\mathcal{ALC}$ TBox* is a finite set of Horn-$\mathcal{ALC}$ CIs. Using the polarity of $\mathcal{ALC}$ concepts, a different definition of Horn concepts is given in [17]. We show in the appendix that both definitions are equivalent. Horn modal formulas are defined in [29]. Again we show in the appendix that, modulo the standard translation between description and modal logic, both definitions are equivalent. In contrast to the Horn modal logic considered in [29], the modal logics called Horn in [26, 8, 10, 7] are not within the scope of this paper: they are not fragments of first-order Horn logic (not even preserved under products) and query evaluation using TBoxes in those languages is coNP-hard (unless additional constraints such as $\forall x \exists y \, r(x, y)$ are imposed on the interpretation of roles).

Interpretations $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ as well as the semantics of $\mathcal{ALC}$ concepts are defined as usual. For example, a CI $C \sqsubseteq D$ *follows from* a TBox $\mathcal{T}$, in symbols $\mathcal{T} \models C \sqsubseteq D$, if every model of $\mathcal{T}$ satisfies $C \sqsubseteq D$. Two concepts C, D are *equivalent* if both $C \sqsubseteq D$ and $D \sqsubseteq C$ follow from the empty TBox. We introduce the notions of simulations between interpretations investigated in this paper, starting with simulations for $\mathcal{EL}$ and $\mathcal{ELU}$. A *pointed interpretation* $(\mathcal{I}, X)$ is an interpretation $\mathcal{I}$ with a set $X \subseteq \Delta^{\mathcal{I}}$. We use $(\mathcal{I}, d)$ to denote $(\mathcal{I}, X)$ if $X = \{d\}$.

Definition 1 (Simulation). *Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. A relation $S \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{J}}$ is a* simulation *between $\mathcal{I}$ and $\mathcal{J}$ if the following conditions hold:*

(A) *if $(d, e) \in S$ and $d \in A^{\mathcal{I}}$, then $e \in A^{\mathcal{J}}$, for all $A \in \mathsf{N_C}$;*
(F) *if $(d, e) \in S$ and $(d, d') \in r^{\mathcal{I}}$, then there exists e' with $(e, e') \in r^{\mathcal{J}}$ and $(d', e') \in S$.*

$(\mathcal{I}, d)$ is simulated by $(\mathcal{J}, e)$, in symbols $(\mathcal{I}, d) \preceq_{sim} (\mathcal{J}, e)$, if there exists a simulation S between $\mathcal{I}$ and $\mathcal{J}$ with $(d, e) \in S$.

Besides the 'infinitary' simulations of Definition 1, we consider simulations with a fixed number of steps.

Definition 2 (k-Simulation). *Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. Define relations $\preceq_{sim}^k$, $k \geq 0$, between pointed interpretations $(\mathcal{I}, d)$ and $(\mathcal{J}, e)$ by induction:*

- *$(\mathcal{I}, d) \preceq_{sim}^0 (\mathcal{J}, e)$ if $d \in A^{\mathcal{I}}$ implies $e \in A^{\mathcal{J}}$, for all $A \in \mathsf{N_C}$.*

  – $(\mathcal{I}, d) \preceq^{k+1}_{sim} (\mathcal{J}, e)$ if $(\mathcal{I}, d) \preceq^{0}_{sim} (\mathcal{J}, e)$ and for all $(d, d') \in r^{\mathcal{I}}$ there exists e' with $(e, e') \in r^{\mathcal{J}}$ and $(\mathcal{I}, d') \preceq^{k}_{sim} (\mathcal{J}, e')$, for all $r \in \mathsf{N_R}$.

We say that $(\mathcal{I}, d)$ is k-simulated by $(\mathcal{J}, e)$ if $(\mathcal{I}, d) \preceq^{k}_{sim} (\mathcal{J}, e)$.

For every $k \geq 0$, we fix a finite set $\mathcal{EL}_k$ of $\mathcal{EL}$ concepts of depth $\leq k$ such that for every $\mathcal{EL}$ concept L of depth $\leq k$ there exists a concept L' in $\mathcal{EL}_k$ such that L and L' are equivalent. For a pointed interpretation $(\mathcal{I}, d)$ we denote by $\lambda_{\mathcal{I},k}(d)$ the conjunction of all concepts L in $\mathcal{EL}_k$ such that $d \in L^{\mathcal{I}}$. The following result linking $\mathcal{EL}$ and $\mathcal{ELU}$ concepts of depth k to k-simulations is folklore and easy to prove.

Lemma 1. *Let $(\mathcal{I}, d)$ and $(\mathcal{J}, e)$ be pointed interpretations. The following conditions are equivalent for all $k \geq 0$:*

1. *$e \in \lambda_{\mathcal{I},k}(d)^{\mathcal{J}}$;*
2. *for all $\mathcal{ELU}$ concepts (equivalently, all $\mathcal{EL}$ concepts) C of depth $\leq k$: if $d \in C^{\mathcal{I}}$, then $e \in C^{\mathcal{J}}$;*
3. *$(\mathcal{I}, d) \preceq^{k}_{sim} (\mathcal{J}, e)$.*

The equivalence of Points (2) and (3) in Lemma 1 can be lifted to arbitrary $\mathcal{ELU}$ concepts and infinitary simulations if $\mathcal{I}$ and $\mathcal{J}$ have finite outdegree or satisfy standard saturatedness conditions [24].

    We now define the new notion of *Horn simulations* which characterizes Horn-$\mathcal{ALC}$ concepts. For a binary relation $\mathcal{R}$ and sets X, Y, we set $X \mathcal{R}^{\uparrow} Y$ if for all $d \in X$ there exists $d' \in Y$ with $(d, d') \in \mathcal{R}$ and we set $X \mathcal{R}^{\downarrow} Y$ if for all $d' \in Y$ there exists $d \in X$ with $(d, d') \in \mathcal{R}$.

Definition 3 (Horn Simulation). *Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. A Horn simulation between $\mathcal{I}$ and $\mathcal{J}$ is a relation $Z \subseteq \mathcal{P}(\Delta^{\mathcal{I}}) \times \Delta^{\mathcal{J}}$ such that if $X \ Z \ d$ then $X \neq \emptyset$ and the following conditions hold:*

(A) *if $X \ Z \ d$ and $X \subseteq A^{\mathcal{I}}$, then $d \in A^{\mathcal{J}}$, for all $A \in \mathsf{N_C}$;*
(F) *if $X \ Z \ d$ and $X(r^{\mathcal{I}})^{\uparrow} Y$, then there exist $Y' \subseteq Y$ and $d' \in \Delta^{\mathcal{J}}$ such that $(d, d') \in r^{\mathcal{J}}$ and $Y' \ Z \ d'$, for all $r \in \mathsf{N_R}$;*
(B) *if $X \ Z \ d$ and $(d, d') \in r^{\mathcal{J}}$, then there exists $Y \subseteq \Delta^{\mathcal{I}}$ with $X(r^{\mathcal{I}})^{\downarrow} Y$ and $Y \ Z \ d'$, for all $r \in \mathsf{N_R}$;*
(S) *$(\mathcal{J}, d) \preceq_{sim} (\mathcal{I}, x)$ for all $x \in X$.*

$(\mathcal{I}, X)$ is Horn simulated by $(\mathcal{J}, d)$, in symbols $(\mathcal{I}, X) \preceq_{horn} (\mathcal{J}, d)$, if there exists a Horn simulation Z between $\mathcal{I}$ and $\mathcal{J}$ such that $X \ Z \ d$.

Example 1. Consider the interpretations $\mathcal{I}_0$ and $\mathcal{J}_0$ depicted in Figure 1. Then the relation

$$Z = \{(\{a, d\}, a'), (\{e\}, b')\}$$

is a Horn simulation between $\mathcal{I}_0$ and $\mathcal{J}_0$.

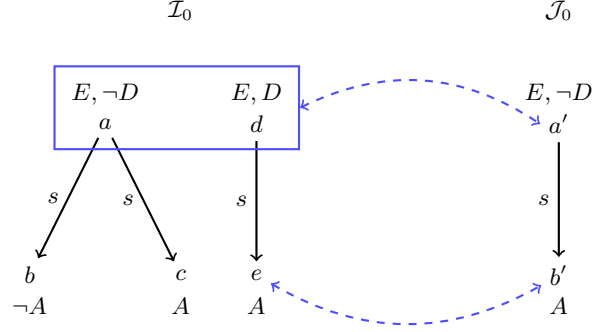We define the finitary approximations of Horn simulations.

Fig. 1. Horn simulation between $\mathcal{I}_0$ and $\mathcal{J}_0$

Definition 4 (k-Horn Simulation). *Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. Define relations $\preceq_{horn}^k$, $k \geq 0$, between pointed interpretations $(\mathcal{I}, X)$ and $(\mathcal{J}, d)$ by induction:*

- *$(\mathcal{I}, X) \preceq_{horn}^0 (\mathcal{J}, d)$ if $X \neq \emptyset$ and $X \subseteq A^{\mathcal{I}}$ implies $d \in A^{\mathcal{J}}$, for all $A \in \mathsf{N_C}$, and $(\mathcal{J}, d) \preceq_{sim}^0 (\mathcal{I}, x)$ for all $x \in X$.*
- *$(\mathcal{I}, X) \preceq_{horn}^{k+1} (\mathcal{J}, d)$ if the following conditions hold:*
  (A) *$(\mathcal{I}, X) \preceq_{horn}^0 (\mathcal{J}, d)$;*
  (F) *if $X(r^{\mathcal{I}})^\uparrow Y$, then there exist $Y' \subseteq Y$ and $d' \in \Delta^{\mathcal{J}}$ such that $(d, d') \in r^{\mathcal{J}}$ and $(\mathcal{I}, Y') \preceq_{horn}^k (\mathcal{J}, d')$, for all $r \in \mathsf{N_R}$;*
  (B) *if $(d, d') \in r^{\mathcal{J}}$, then there exists $Y \subseteq \Delta^{\mathcal{I}}$ with $X(r^{\mathcal{I}})^\downarrow Y$ and $(\mathcal{I}, Y) \preceq_{horn}^k (\mathcal{J}, d')$, for all $r \in \mathsf{N_R}$;*
  (S) *$(\mathcal{J}, d) \preceq_{sim}^{k+1} (\mathcal{I}, x)$ for all $x \in X$.*

We say that $(\mathcal{I}, X)$ is k-Horn-simulated by $(\mathcal{J}, d)$ if $(\mathcal{I}, X) \preceq_{horn}^k (\mathcal{J}, d)$.

For every $k \geq 0$, we fix a finite set Horn_k of Horn-$\mathcal{ALC}$ concepts of depth $\leq k$ such that for every Horn-$\mathcal{ALC}$ concept R of depth $\leq k$ there exists a concept R' in Horn_k such that R and R' are equivalent. For a pointed interpretation $(\mathcal{I}, X)$ we denote by $\rho_{\mathcal{I},k}(X)$ the conjunction of all concepts R in Horn_k such that $X \subseteq R^{\mathcal{I}}$. We establish a close link between k-Horn simulations and Horn-$\mathcal{ALC}$ concepts of depth $\leq k$.

Lemma 2. *Let $(\mathcal{I}, X)$ and $(\mathcal{J}, d)$ be pointed interpretations. Then the following conditions are equivalent, for all $k \geq 0$:*

1. *$d \in \rho_{\mathcal{I},k}(X)^{\mathcal{J}}$;*
2. *for all Horn-$\mathcal{ALC}$ concepts R of depth $\leq k$: if $X \subseteq R^{\mathcal{I}}$, then $d \in R^{\mathcal{J}}$;*
3. *there exists a set $X_0 \subseteq X$ such that $(\mathcal{I}, X_0) \preceq_{horn}^k (\mathcal{J}, d)$.*

Proof. (1.) $\Leftrightarrow$ (2.) holds by definition of $\rho_{\mathcal{I},k}(X)$. To show (2.) $\Rightarrow$ (3.), define relations $Z_k \subseteq \mathcal{P}(\Delta^{\mathcal{I}}) \times \Delta^{\mathcal{J}}$, $k \geq 0$, by setting $(Y, e) \in Z_k$ if $Y \neq \emptyset$ and

- $e \in \rho_{\mathcal{I},k}(Y)^{\mathcal{J}}$;

– $Y \subseteq \lambda_{\mathcal{J},k}(e)^{\mathcal{I}}$.

We show the following claim by induction over $k \geq 0$.

Claim 1. For all $k \geq 0$, $Y \subseteq \Delta^{\mathcal{I}}$ and $e \in \Delta^{\mathcal{J}}$, if $Y\ Z_k\ e$, then $(\mathcal{I}, Y) \preceq^k_{\text{horn}} (\mathcal{J}, e)$.

For $k = 0$, Claim 1 holds by definition. Assume Claim 1 has been proved for k and let $Y\ Z_{k+1}\ e$. We show $(\mathcal{I}, Y) \preceq^{k+1}_{\text{horn}} (\mathcal{J}, e)$. Point (A) holds by definition. For Point (F), assume $Y(r^{\mathcal{I}})^{\uparrow}V$. We then have $Y \subseteq (\exists r.\rho_{\mathcal{I},k}(V))^{\mathcal{I}}$. By definition of Z_{k+1}, $e \in (\exists r.\rho_{\mathcal{I},k}(V))^{\mathcal{J}}$. Thus, there exists e' with $(e, e') \in r^{\mathcal{J}}$ and $e' \in \rho_{\mathcal{I},k}(V)^{\mathcal{J}}$. Set

$$Y' = V \cap \lambda_{\mathcal{J},k}(e')^{\mathcal{I}}.$$

We show that (Y', e') is as required for Point (F). By induction hypothesis, it suffices to prove that $Y'\ Z_k\ e'$:

– We show that $Y' \neq \emptyset$. Assume $Y' = \emptyset$. Then $V \subseteq (\lambda_{\mathcal{J},k}(e') \to \bot)^{\mathcal{I}}$. Then $(\lambda_{\mathcal{J},k}(e') \to \bot)$ is equivalent to a conjunct of $\rho_{\mathcal{I},k}(V)$. Then $e' \in (\lambda_{\mathcal{J},k}(e') \to \bot)^{\mathcal{J}}$, by construction of e'. We also have $e' \in \lambda_{\mathcal{J},k}(e')^{\mathcal{J}}$ which is a contradiction.
– By definition, $Y' \subseteq \lambda_{\mathcal{J},k}(e')^{\mathcal{I}}$.
– Assume $Y' \subseteq R^{\mathcal{I}}$ for some concept R in Horn_k. We have to show that $e' \in R^{\mathcal{J}}$. From $Y' \subseteq R^{\mathcal{I}}$ and the definition of Y' we obtain $V \subseteq (\lambda_{\mathcal{J},k}(e') \to R)^{\mathcal{I}}$. Then $(\lambda_{\mathcal{J},k}(e') \to R)$ is equivalent to a conjunct of $\rho_{\mathcal{I},k}(V)$. By construction of e', $e' \in (\lambda_{\mathcal{J},k}(e') \to R)^{\mathcal{J}}$. We also have $e' \in \lambda_{\mathcal{J},k}(e')^{\mathcal{J}}$. Thus $e' \in R^{\mathcal{J}}$, as required.

To show Point (B), assume that $(e, e') \in r^{\mathcal{J}}$. Take for every concept R in Horn_k with $e' \notin R^{\mathcal{J}}$ some $e_R \in Y$ and e'_R with $(e_R, e'_R) \in r^{\mathcal{I}}$ such that $e'_R \in (\lambda_{\mathcal{J},k}(e') \sqcap \neg R)^{\mathcal{I}}$. They exist because otherwise $Y \subseteq (\forall r.(\lambda_{\mathcal{J},k}(e') \to R))^{\mathcal{I}}$ but $e \notin (\forall r.(\lambda_{\mathcal{J},k}(e') \to R))^{\mathcal{J}}$ which contradicts the definition of Z_{k+1}. Let Y' be the set of all such e'_R. Then $Y(r^{\mathcal{I}})^{\downarrow}Y'$ and $Y'\ Z_k\ e'$ as $e' \in \rho_{\mathcal{I},k}(Y')^{\mathcal{J}}$ and $Y' \subseteq \lambda_{\mathcal{J},k}(e')^{\mathcal{I}}$ hold by construction of Y'. By induction hypothesis, $(\mathcal{I}, Y') \preceq^k_{\text{horn}} (\mathcal{J}, e')$ and Point (B) follows.

Point (S) follows from Lemma 1 and so Claim 1 is proved.

It remains to prove that if $d \in \rho_{\mathcal{I},k}(X)^{\mathcal{J}}$, then there exists $X_0 \subseteq X$ with $(\mathcal{I}, X_0) \preceq^k_{\text{horn}} (\mathcal{J}, d)$. Set $X_0 = X \cap \lambda_{\mathcal{J},k}(d)^{\mathcal{I}}$. Then $(\mathcal{I}, X_0) \preceq^k_{\text{horn}} (\mathcal{J}, d)$ can be proved in the same way as Claim 1 above.

For the proof of (3.) $\Rightarrow$ (2.), it suffices to prove the following claim by induction over $k \geq 0$.

Claim 2. For all $k \geq 0$, $Y \subseteq \Delta^{\mathcal{I}}$ and $e \in \Delta^{\mathcal{J}}$, if $(\mathcal{I}, Y) \preceq^k_{\text{horn}} (\mathcal{J}, e)$, then $Y \subseteq R^{\mathcal{I}}$ implies $e \in R^{\mathcal{I}}$ for all R in Horn_k.

For $k = 0$, Claim 2 follows by definition. Assume Claim 2 has been proved for k. We prove Claim 2 for $k + 1$ by induction over the construction of R. Thus, assume that Claim 2 has been proved for $R', R_1, R_2 \in \text{Horn}_{k+1}$, and that $R \in \text{Horn}_{k+1}$ is of the form $R = \forall r.R'$, $R = \exists r.R'$, $R = R_1 \sqcap R_2$, or $R = L \to R'$. Then we prove Claim 2 for R. Assume $(\mathcal{I}, Y) \preceq^{k+1}_{\text{horn}} (\mathcal{J}, e)$ and $Y \subseteq R^{\mathcal{I}}$.

- Assume $R = \forall r.R'$ and for a proof by contradiction that $e \notin (\forall r.R')^{\mathcal{J}}$. Choose e' with $(e, e') \in r^{\mathcal{J}}$ and $e' \notin R'^{\mathcal{J}}$. By (B), there exists $Y' \subseteq \Delta^{\mathcal{I}}$ with $Y(r^{\mathcal{I}})^{\downarrow}Y'$ and $(\mathcal{I}, Y') \preceq_{\text{horn}}^k (\mathcal{J}, e')$. We have $R' \in \text{Horn}_k$ and thus, as we assume that Claim 2 has been proved for k, there exists $e'' \in Y'$ with $e'' \notin R'^{\mathcal{I}}$. Then there exists $e''' \in Y$ with $e''' \notin (\forall r.R')^{\mathcal{I}}$, and we have derived a contradiction.
- Assume $R = \exists r.R'$ and for a proof by contradiction that $e \notin (\exists r.R')^{\mathcal{J}}$. We have $Y(r^{\mathcal{I}})^{\uparrow}R'^{\mathcal{I}}$. Thus, by (F), there exist $Y' \subseteq R'^{\mathcal{I}}$ and $e' \in \Delta^{\mathcal{J}}$ with $(e, e') \in r^{\mathcal{J}}$ and $(\mathcal{I}, Y') \preceq_{\text{horn}}^k (\mathcal{J}, e')$. We have $R' \in \text{Horn}_k$ and thus, as we assume that Claim 2 has been proved for k, $e' \in R'^{\mathcal{J}}$. But then $e \in (\exists r.R')^{\mathcal{J}}$. We have derived a contradiction.
- Assume $R = R_1 \sqcap R_2$ and for a proof by contradiction that $e \notin (R_1 \sqcap R_2)^{\mathcal{J}}$. We may assume w.l.o.g. that $e \notin R_1^{\mathcal{J}}$. But then from $Y \subseteq R^{\mathcal{I}}$ we obtain $Y \subseteq R_1^{\mathcal{I}}$ and we have derived a contradiction to the induction hypothesis.
- Assume $R = (L \to R')$ and for a proof by contradiction $e \notin (L \to R')^{\mathcal{J}}$. Then $e \in L^{\mathcal{J}}$ and $e \notin R'^{\mathcal{J}}$. By Lemma 1, $Y \subseteq L^{\mathcal{I}}$, and, by induction hypothesis, there exists $e' \in Y$ with $e' \notin R'^{\mathcal{I}}$. Then $Y \nsubseteq (L \to R')^{\mathcal{I}}$ and we have derived a contradiction.

This finishes the proof. □

We will also be using bisimulations in some proofs. Recall that a relation S between $\mathcal{I}$ and $\mathcal{J}$ is a *bisimulation* if it is a simulation between $\mathcal{I}$ and $\mathcal{J}$ and its inverse is a simulation between $\mathcal{J}$ and $\mathcal{I}$. *k-bisimulations* are defined accordingly. Then pointed interpretations $(\mathcal{I}, d)$ and $(\mathcal{I}, e)$ are k-bisimilar if, and only if, $d \in C^{\mathcal{I}}$ iff $e \in C^{\mathcal{J}}$ holds for all $\mathcal{ALC}$ concepts C of depth $\leq k$. Van Benthem's Theorem (extended by depth information) can be formulated as follows: a FO-sentence $\varphi(x)$ of quantifier depth k is invariant under bisimulations iff it is invariant under $(2^k - 1)$-bisimulations iff it is equivalent to a $\mathcal{ALC}$-concept [28].

The above notions of simulations, bisimulations, and Horn simulations can be captured by games similar to the Ehrenfeucht Fraïssé games in the obvious way. We do not spell out the details here but note that for Horn simulations the players have to select *subsets* rather than elements of the interpretation.

3 Characterization of Horn-$\mathcal{ALC}$ Concepts

In this section, we prove the characterization of Horn-$\mathcal{ALC}$ concepts via preservation under Horn simulations. We say that an $\mathcal{ALC}$ concept C is *preserved under (k-)Horn simulations* if for all pointed interpretations $(\mathcal{I}, X)$ and $(\mathcal{J}, d)$, $X \subseteq C^{\mathcal{I}}$ and $(\mathcal{I}, X) \preceq_{\text{horn}}^{(k)} (\mathcal{J}, d)$ imply $d \in C^{\mathcal{J}}$.

Theorem 1. *Let C be an $\mathcal{ALC}$ concept of depth k. Then the following conditions are equivalent:*

1. *C is equivalent to a Horn-$\mathcal{ALC}$ concept;*
2. *C is preserved under Horn simulations;*

3. C is preserved under k-Horn simulations.

(1.) $\Rightarrow$ (3.) follows from Lemma 2 and (3.) $\Rightarrow$ (2.) is trivial. It thus remains to prove (2.) $\Rightarrow$ (1.). We prove this implication in two steps. Let C be an $\mathcal{ALC}$ concept of depth $\leq k$. Then we first show that preservation under k-Horn simulations entails equivalence to a Horn-$\mathcal{ALC}$ concept of depth k, and then we show that preservation under Horn simulations already entails preservation under k-Horn simulations.

Lemma 3. *If C is an $\mathcal{ALC}$ concept of depth $\leq k$ preserved under k-Horn simulations, then C is equivalent to a Horn-$\mathcal{ALC}$ concept of depth $\leq k$.*

Proof. Let C be an $\mathcal{ALC}$ concept of depth $\leq k$ preserved under k-Horn simulations. Let D be the conjunction of all $R \in \mathrm{Horn}_k$ with $\emptyset \models C \sqsubseteq R$. It suffices to show that $\emptyset \models D \sqsubseteq C$ because then D and C are equivalent. To show this, observe that D is equivalent to $\rho_{\mathcal{I}_0,k}(X)$, where $\mathcal{I}_0$ is the disjoint union of all finite interpretations (up to isomorphisms) and $X = C^{\mathcal{I}_0}$. (To prove this, assume first that $R \in \mathrm{Horn}_k$ and $\emptyset \models C \sqsubseteq R$. Then $X = C^{\mathcal{I}} \subseteq R^{\mathcal{I}}$ for all interpretations $\mathcal{I}$. Thus, R is a conjunct of $\rho_{\mathcal{I}_0,k}(X)$. Conversely, assume that R is a conjunct of $\rho_{\mathcal{I}_0,k}(X)$. Then $X \subseteq R^{\mathcal{I}_0}$. Thus, $C^{\mathcal{I}_0} \subseteq R^{\mathcal{I}_0}$. But then $\emptyset \models C \sqsubseteq R$ since the latter is equivalent to the condition that $C^{\mathcal{I}} \subseteq R^{\mathcal{I}}$ for all *finite* interpretations $\mathcal{I}$ and $\mathcal{ALC}$ has the finite model property.) Now assume that $(\mathcal{J}, d)$ is a pointed interpretation with $d \in D^{\mathcal{J}}$. Then $d \in \rho_{\mathcal{I}_0,k}(X)^{\mathcal{J}}$. Thus, by Lemma 2, there exists $X_0 \subseteq X$ such that $(\mathcal{I}_0, X_0) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}, d)$. Then $d \in C^{\mathcal{J}}$ follows from the assumption that C is preserved under k-Horn simulations. $\square$

An interpretation $\mathcal{I}$ is *tree-shaped* if $G_{\mathcal{I}} = (\Delta^{\mathcal{I}}, E)$ with $E = \bigcup_{r \in \mathsf{N_R}} r^{\mathcal{I}}$ is a directed tree and $r^{\mathcal{I}} \cap s^{\mathcal{I}} = \emptyset$ for all role names $r \neq s$. The root of $G_{\mathcal{I}}$ is called the *root* of $\mathcal{I}$. The *depth* of $d \in \Delta^{\mathcal{I}}$ is the length of the shortest path from d to the root of $\mathcal{I}$. Thus, the root of $\mathcal{I}$ has depth 0. The disjoint union of tree-shaped interpretations is a *forest*. Recall that every pointed $(\mathcal{I}, d)$ can be unravelled into a tree-shaped interpretation $\mathcal{J}$ with root d such that $(\mathcal{I}, d)$ and $(\mathcal{J}, d)$ are bisimilar.

Lemma 4. *Assume C is an $\mathcal{ALC}$ concept of depth k preserved under Horn simulations. Then C is preserved under k-Horn simulations.*

Proof. Assume C is an $\mathcal{ALC}$ concept of depth $\leq k$ preserved under k-Horn simulations. Let $(\mathcal{I}, X)$ and $(\mathcal{J}, d)$ be pointed interpretations such that $(\mathcal{I}, X) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}, d)$ and $X \subseteq C^{\mathcal{I}}$. We have to show that $d \in C^{\mathcal{J}}$. Take for every $x \in X$ a tree-shaped pointed interpretation $(\mathcal{I}_x, x)$ bisimilar to $(\mathcal{I}, x)$. Let $(\mathcal{J}', d)$ be a tree-shaped pointed interpretation bisimilar to $(\mathcal{J}, d)$. Then $(\mathcal{I}', X) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}', d)$ for the disjoint union $\mathcal{I}'$ of $\mathcal{I}_x$, $x \in X$. By bisimulation invariance of $\mathcal{ALC}$ concepts, we have $X \subseteq C^{\mathcal{I}'}$ and it suffices to prove that $d \in C^{\mathcal{J}'}$. Remove from $\mathcal{I}'$ and $\mathcal{J}'$ all nodes of depth $> k$ and denote the resulting interpretations by $\mathcal{I}''$ and $\mathcal{J}''$, respectively. As C has depth $\leq k$, we have $X \subseteq C^{\mathcal{I}''}$ and it suffices to prove that $d \in C^{\mathcal{J}''}$. Using $(\mathcal{I}', X) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}', d)$, it is straightforward to prove that $(\mathcal{I}'', X) \preceq_{\mathrm{horn}} (\mathcal{J}'', d)$. Then $d \in C^{\mathcal{J}''}$ follows from preservation of C under Horn simulations. $\square$

Recall that a *basic Horn FO formula* is a disjunction $\phi_1 \vee \cdots \vee \phi_n$ of first-order logic formulas in which at most one is an atomic formula, the rest being negations of atomic formulas. A *Horn FO formula* is constructed from basic Horn formulas using the connectives $\wedge$, $\exists$, and $\forall$. Horn FO formulas are characterized as the fragment of FO that is preserved under reduced products [9].

Example 2. The $\mathcal{ALC}$ concept $C = ((\exists s.\top) \sqcap ((E \sqcap \forall s.A) \rightarrow D))$ is not preserved under Horn simulations. To show this, observe that, for the interpretations $\mathcal{I}_0$ and $\mathcal{J}_0$ defined in Example 1, $\{a, d\} \subseteq C^{\mathcal{I}_0}$ but $a' \notin C^{\mathcal{J}_0}$. Thus, C is not equivalent to any Horn-$\mathcal{ALC}$ concept. C is, however, equivalent to the Horn FO formula $\exists y\, (s(x, y) \wedge (\neg E(x) \vee \neg A(y) \vee D(x)))$.

4 Characterization of Horn-$\mathcal{ALC}$ TBoxes

In this section, we prove the characterization of Horn-$\mathcal{ALC}$ TBoxes via preservation under global Horn simulations. An $\mathcal{ALC}$ TBox is *preserved under global (k-)Horn simulations* if an interpretation $\mathcal{J}$ is a model of $\mathcal{T}$ whenever for every $d \in \Delta^{\mathcal{J}}$ there exists a model $\mathcal{I}$ of $\mathcal{T}$ and $X \subseteq \Delta^{\mathcal{I}}$ such that $(\mathcal{I}, X) \preceq_{\text{horn}}^{(k)} (\mathcal{J}, d)$.

Theorem 2. *Let $\mathcal{T}$ be an $\mathcal{ALC}$ TBox of depth k. Then the following conditions are equivalent:*

1. *$\mathcal{T}$ is equivalent to a Horn-$\mathcal{ALC}$ TBox;*
2. *$\mathcal{T}$ is preserved under global Horn simulations;*
3. *$\mathcal{T}$ is preserved under global k-Horn simulations.*

(1.) $\Rightarrow$ (3.) follows from Lemma 2 and (3.) $\Rightarrow$ (2.) is trivial. It thus remains to prove (2.) $\Rightarrow$ (1.). As in the concept case, we prove this implication in two steps. Let $\mathcal{T}$ be an $\mathcal{ALC}$ concept of depth $\leq k$. Then we first show that preservation under global k-Horn simulations entails equivalence to a Horn-$\mathcal{ALC}$ TBox of depth $\leq k$, and then we show that preservation under global Horn simulations already entails preservation under global k-Horn simulations.

Lemma 5. *Let $\mathcal{T}$ be an $\mathcal{ALC}$ TBox of depth k preserved under global k-Horn simulations. Then $\mathcal{T}$ is equivalent to a Horn-$\mathcal{ALC}$ TBox of depth $\leq k$.*

Proof. Let $\mathcal{T}'$ be the set of Horn-$\mathcal{ALC}$ CIs $\top \sqsubseteq R$ with R of depth $\leq k$ such that $\mathcal{T} \models \top \sqsubseteq R$. We show that $\mathcal{T}' \models \mathcal{T}$. Take a model $\mathcal{J}$ of $\mathcal{T}'$. Take for every Horn-$\mathcal{ALC}$ CI $\top \sqsubseteq R$ of depth $\leq k$ and $e \in \Delta^{\mathcal{J}} \setminus R^{\mathcal{J}}$ a model $\mathcal{I}_{e,R}$ of $\mathcal{T}$ and $x_{e,R} \in \Delta^{\mathcal{I}_{e,R}}$ with $x_{e,R} \notin R^{\mathcal{I}_{e,R}}$. Such a model exists since $\mathcal{T}' \not\models \top \rightarrow R$ and so, by the definition of $\mathcal{T}'$, $\mathcal{T} \not\models \top \rightarrow R$.

Let $\mathcal{I}$ be the disjoint union of all $\mathcal{I}_{e,R}$ and let X_e be the set of all $x_{e,R}$. Then $X_e \subseteq R^{\mathcal{I}}$ implies $e \in R^{\mathcal{J}}$, for every Horn-$\mathcal{ALC}$ concept R of depth $\leq k$. Thus, by Lemma 2, there exists $Y_e \subseteq X_e$ such that $(\mathcal{I}, Y_e) \preceq_{\text{horn}}^k (\mathcal{J}, e)$. Then $\mathcal{J}$ is model of $\mathcal{T}$ since $\mathcal{I}$ is a model of $\mathcal{T}$ and $\mathcal{T}$ is preserved under global k-Horn simulations. $\qquad\square$

To proceed, we need a lemma connecting Horn simulations and products. Consider a family $\mathcal{I}_i$, $i \in I$, of interpretations. The product

$$\prod_{i \in I} \mathcal{I}_i$$

is defined in the standard way: the domain $\Delta^{\prod_{i \in I} \mathcal{I}_i}$ of $\prod_{i \in I} \mathcal{I}_i$ is the set of functions $f : I \to \bigcup_{i \in I} \Delta^{\mathcal{I}_i}$ such that $f(i) \in \Delta^{\mathcal{I}_i}$ for all $i \in I$. Then

- $A^{\prod_{i \in I} \mathcal{I}_i} = \{f \in \Delta^{\prod_{i \in I} \mathcal{I}_i} \mid \forall i \in I : f(i) \in A^{\mathcal{I}_i}\}$
- $r^{\prod_{i \in I} \mathcal{I}_i} = \{(f,g) \in (\Delta^{\prod_{i \in I} \mathcal{I}_i})^2 \mid \forall i \in I : (f(i), g(i)) \in r^{\mathcal{I}_i}\}$

Now assume that $\mathcal{I}$ is the disjoint union of interpretations $\mathcal{I}_i, i \in I$. Define a relation Z between $\mathcal{P}(\Delta^{\mathcal{I}})$ and $\mathcal{J} = \prod_{i \in I} \mathcal{I}_i$ by setting $(Y, f) \in Z$ if $Y \subseteq \Delta^{\mathcal{I}}$, $f \in \Delta^{\mathcal{J}}$, and $\Delta^{\mathcal{I}_i} \cap Y = \{f(i)\}$ for all $i \in I$. Note that it follows in particular that for all Y with $(Y, f) \in Z$ for some f, Y contains exactly one node from each $\Delta^{\mathcal{I}_i}$, for $i \in I$.

Lemma 6. *Z is a Horn simulation between $\mathcal{I}$ and $\mathcal{J}$.*

Proof. Conditions (A) and (S) are straightforward [23].

For Condition (F), assume that $(X, f) \in Z$ and that $X(r^{\mathcal{I}})^{\uparrow}Y$. Let $Y_0 \subseteq Y$ contain for every $f(i) \in X$ exactly one $g(i) \in Y$ with $(f(i), g(i)) \in r^{\mathcal{I}_i}$. Then $(Y_0, g) \in Z$, as required.

For Condition (B), assume that $(X, f) \in Z$ and that $(f, g) \in r^{\mathcal{J}}$. Let $Y = \{g(i) \mid i \in I\}$. Then $X(r^{\mathcal{I}})^{\downarrow}Y$ and $(Y, g) \in Z$, as required. $\square$

We require the following observation. Let $\mathcal{I}$ and $\mathcal{J}$ be interpretations. Call a sequence $H^0, \ldots, H^k$ of relations a *k-Horn simulation* if $(X, d) \in H^i$ implies $(\mathcal{I}, X) \preceq^i_{\text{horn}} (\mathcal{J}, d)$, for all $0 \leq i \leq k$, and the following holds for $0 \leq i < k$ and all $(X, d) \in H^{i+1}$:

- if $X(r^{\mathcal{I}})^{\uparrow}Y$, then there exist $Y' \subseteq Y$ and $d' \in \Delta^{\mathcal{J}}$ such that $(d, d') \in r^{\mathcal{J}}$ and $(Y', d') \in H^i$, for all $r \in \mathsf{N_R}$;
- if $(d, d') \in r^{\mathcal{J}}$, then there exists $Y \subseteq \Delta^{\mathcal{I}}$ with $X(r^{\mathcal{I}})^{\downarrow}Y$ and $(Y', d') \in H^i$, for all $r \in \mathsf{N_R}$.

Let $(\mathcal{I}, X) \preceq^k_{\text{horn}} (\mathcal{J}, d)$. Take for every $x \in X$ a tree-shaped pointed interpretation $(\mathcal{I}_x, x)$ bisimilar to $(\mathcal{I}, x)$. Let $(\mathcal{J}', d)$ be a tree-shaped pointed interpretation bisimilar to $(\mathcal{J}, d)$. Then $(\mathcal{I}', X) \preceq^k_{\text{horn}} (\mathcal{J}', d)$ for the disjoint union $\mathcal{I}'$ of the interpretations $\mathcal{I}_x, x \in X$. By duplicating successors in $\mathcal{J}'$ sufficiently often (possibly exponentially many times) we obtain a tree-shaped pointed interpretation $(\mathcal{J}'', d)$ bisimilar to $(\mathcal{J}', d)$ such that there is a k-Horn simulation $H^0, \ldots, H^k$ between $\mathcal{I}'$ and $\mathcal{J}''$ with

- $(X, d) \in H^k$;
- for all $0 \leq i \leq k$, if $(X_0, e), (X_1, e) \in H^i$, then $X_0 = X_1$.

Lemma 7. *Suppose $\mathcal{T}$ has depth k and is preserved under global Horn simulations. Then $\mathcal{T}$ is preserved under global k-Horn simulations.*

Proof. Let $\mathcal{J}$ be an interpretation such that for every $d \in \Delta^{\mathcal{J}}$ there exists a model $\mathcal{I}$ of $\mathcal{T}$ and $X \subseteq \Delta^{\mathcal{I}}$ with $(\mathcal{I}, X) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}, d)$. We have to show that $\mathcal{J}$ is a model of $\mathcal{T}$. Let $d \in \Delta^{\mathcal{J}}$ and take a model $\mathcal{I}$ of $\mathcal{T}$ and $X \subseteq \Delta^{\mathcal{I}}$ with $(\mathcal{I}, X) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}, d)$. It suffices to show that $d \in (\neg C \sqcup D)^{\mathcal{J}}$ for all $C \sqsubseteq D \in \mathcal{T}$. Take for every $x \in X$ a tree-shaped pointed interpretation $(\mathcal{I}_x, x)$ bisimilar to $(\mathcal{I}, x)$. By the observation above, we can take a tree-shaped pointed interpretation $(\mathcal{J}', d)$ bisimilar to $(\mathcal{J}, d)$ such that $(\mathcal{I}', X) \preceq_{\mathrm{horn}}^{k} (\mathcal{J}', d)$ for the disjoint union $\mathcal{I}'$ of the interpretations $\mathcal{I}_x$, $x \in X$, and, moreover, there is a k-Horn simulation $H^0, \ldots, H^k$ between $\mathcal{I}'$ and $\mathcal{J}'$ with

- $(X, d) \in H^k$;
- for all $0 \leq i \leq k$, if $(X_0, e), (X_1, e) \in H^i$, then $X_0 = X_1$.

By bisimulation invariance of $\mathcal{ALC}$ concepts, $\mathcal{I}'$ is a model of $\mathcal{T}$ and it suffices to prove that $d' \in (\neg C \sqcup D)^{\mathcal{J}'}$ for all $C \sqsubseteq D \in \mathcal{T}$.

To prove this, we do the following: we consider the interpretation $\mathcal{J}'_{|k}$ obtained from $\mathcal{J}'$ by removing all nodes of depth $> k$. Then we hook to every leaf $e \in \mathcal{J}'_{|k}$ of depth k a tree-shaped interpretation $\mathcal{I}_e$ in such a way that that $(\mathcal{I}', X) \preceq_{\mathrm{horn}} (\mathcal{J}'', d')$ for the resulting interpretation $\mathcal{J}''$. By preservation of $\mathcal{T}$ under Horn simulations, $\mathcal{J}''$ is a model of $\mathcal{T}$. As $\mathcal{T}$ has depth $\leq k$, $d' \in (\neg C \sqcup D)^{\mathcal{J}'}$ for all $C \sqsubseteq D \in \mathcal{T}$. We come to the construction of the $\mathcal{I}_e$. Assume e of depth k in $\mathcal{J}'$ is given. Then there is a unique nonempty $X \subseteq \Delta^{\mathcal{I}'}$ such that

- all $x \in X$ have depth k in $\mathcal{I}'$;
- $(X, e) \in H^0$.

Observe that Point (2) implies that for every $A \in \mathsf{N_C}$: $X \subseteq A^{\mathcal{I}'}$ iff $e \in A^{\mathcal{J}'}$. Denote, for any $x \in \Delta^{\mathcal{I}'}$, by $\mathcal{I}'_x$ the tree-shaped subinterpretation of $\mathcal{I}'$ rooted at x. Now we hook to e the interpretation $\prod_{x \in X} \mathcal{I}'_x$ by identifying $(x \mid x \in X) \in \prod_{x \in X} \mathcal{I}'_x$ with e. Using Lemma 6, it is readily checked that the resulting interpretation is as required. $\qquad\square$

We now give an example of an $\mathcal{ALC}$ TBox not preserved under global Horn simulations which is equivalent to a Horn FO sentence. Note that such TBoxes are still rather well behaved in OBDA applications: it is shown in [15] that conjunctive query evaluation is in PTime for any $\mathcal{ALC}$ TBox preserved under products.

Example 3. The $\mathcal{ALC}$ TBox $\mathcal{T} = \{E \sqsubseteq \exists s.\top, E \sqcap \forall s.A \sqsubseteq D\}$ is not preserved under global Horn simulations. To show this, observe that, for the interpretations $\mathcal{I}_0$ and $\mathcal{J}_0$ and Horn simulation Z defined in Example 1, $\mathcal{I}_0$ is a model of $\mathcal{T}$ but $\mathcal{J}_0$ is not. Thus, $\mathcal{T}$ is not equivalent to any Horn-$\mathcal{ALC}$ TBox. $\mathcal{T}$ is, however, equivalent to the Horn FO-sentence

$$\forall x \exists y ((\neg E(x) \lor s(x, y)) \land (\neg E(x) \lor \neg A(y) \lor D(x))).$$

5 Conclusion

We have given a model-theoretic characterization of the expressive power of Horn-$\mathcal{ALC}$ concepts and TBoxes via Horn simulations. Weaker Horn DLs such as $\mathcal{EL}$ and DL-Lite have already been characterized in [23]. By introducing variations of Horn simulations, many model-theoretic characterization problems can now be attacked: we conjecture that the obvious two-way Horn simulations characterise Horn-$\mathcal{ALCI}$ concepts and TBoxes and that 'graded' Horn simulations can be used to characterize Horn-$\mathcal{ALCIQ}$. Using techniques from [11], it should also be possible to characterize Horn-$\mathcal{SHIQ}$. It would also be of interest to consider Horn fragments of the guarded fragment or other important fragments of FO and to confirm that the characterization results still hold in the sense of finite model theory, for example: an $\mathcal{ALC}$ concept is preserved under Horn simulations between *finite* interpretations iff it is equivalent to a Horn-$\mathcal{ALC}$ concept.

Acknowledgements

References

1. Baader, F., Calvanese, D., McGuiness, D., Nardi, D., Patel-Schneider, P.: The Description Logic Handbook: Theory, implementation and applications. Cambridge University Press (2003)
2. Baader, F., Bienvenu, M., Lutz, C., Wolter, F.: Query and predicate emptiness in ontology-based data access. J. Artif. Intell. Res. (JAIR) 56, 1–59 (2016)
3. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: An Introduction to Description Logic. Cambridge University Press (2017)
4. van Benthem, J.: Modal Logic and Classical Logic. Bibliopolis (1983)
5. Bienvenu, M., Hansen, P., Lutz, C., Wolter, F.: First order-rewritability and containment of conjunctive queries in horn description logics. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016. pp. 965–971 (2016)
6. Botoeva, E., Kontchakov, R., Ryzhikov, V., Wolter, F., Zakharyaschev, M.: Games for query inseparability of description logic knowledge bases. Artif. Intell. 234, 78–119 (2016)
7. Bresolin, D., Muñoz-Velasco, E., Sciavicco, G.: On the expressive power of subpropositional fragments of modal logic. In: Proc. GandALF16. pp. 91–104 (2016)
8. Fariñas del Cerro, L., Penttonen, M.: A note of the complexity of the satisfiability of modal horn clauses. J. Log. Program. 4(1), 1–10 (1987)
9. Chang, C.C., Keisler, H.J.: Model Theory. Elsevier (1990)
10. Chen, C.C., Lin, I.P.: The computational complexity of satisfiability of temporal horn formulas in propositional linear-time temporal logic. Information Processing Letters 45(3), 131–136 (1993)
11. Dawar, A., Otto, M.: Modal characterisation theorems over special classes of frames. Ann. Pure Appl. Logic 161(1), 1–42 (2009)

12. Eiter, T., Gottlob, G., Ortiz, M., Simkus, M.: Query answering in the description logic horn-$\mathcal{SHIQ}$. In: Logics in Artificial Intelligence, 11th European Conference, JELIA 2008, Dresden, Germany, September 28 - October 1, 2008. Proceedings. pp. 166–179 (2008)

13. Glimm, B., Kazakov, Y., Tran, T.: Ontology materialization by abstraction refinement in horn SHOIF. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA. pp. 1114–1120 (2017)

14. Gutiérrez-Basulto, V., Jung, J.C., Sabellek, L.: Reverse engineering queries in ontology-enriched systems: The case of expressive horn description logic ontologies. In: Proceedings of IJCAI-ECAI-18. AAAI Press (2018)

15. Hernich, A., Lutz, C., Papacchini, F., Wolter, F.: Dichotomies in ontology-mediated querying with the guarded fragment (2017), journal version of PODS'2017 paper

16. Hernich, A., Lutz, C., Papacchini, F., Wolter, F.: Horn-rewritability vs ptime query evaluation in ontology-mediated querying. In: Proceedings of IJCAI-ECAI. AAAI Press (2018)

17. Hustadt, U., Motik, B., Sattler, U.: Data complexity of reasoning in very expressive description logics. In: IJCAI. pp. 466–471 (2005)

18. Hustadt, U., Motik, B., Sattler, U.: Reasoning in description logics by a reduction to disjunctive datalog. J. Autom. Reasoning 39(3), 351–384 (2007)

19. Kazakov, Y.: Consequence-driven reasoning for Horn-$\mathcal{SHIQ}$ ontologies. In: Boutilier, C. (ed.) IJCAI. pp. 2040–2045 (2009)

20. Krötzsch, M.: Description Logic Rules, Studies on the Semantic Web, vol. 8. IOS Press (2010), https://doi.org/10.3233/978-1-61499-342-1-i

21. Krötzsch, M., Rudolph, S., Hitzler, P.: Complexities of horn description logics. ACM Trans. Comput. Log. 14(1), 2:1–2:36 (2013), http://doi.acm.org/10.1145/2422085.2422087

22. Kurtonina, N., de Rijke, M.: Expressiveness of concept expressions in first-order description logics. Artif. Intell. 107(2), 303–333 (1999)

23. Lutz, C., Piro, R., Wolter, F.: Description logic TBoxes: Model-theoretic characterizations and rewritability. In: IJCAI (2011)

24. Lutz, C., Wolter, F.: Deciding inseparability and conservative extensions in the description logic EL. J. Symb. Comput. 45(2), 194–228 (2010)

25. Lutz, C., Wolter, F.: The data complexity of description logic ontologies. Logical Methods in Computer Science 13(4) (2017)

26. Nguyen, L.A.: On the complexity of fragments of modal logics. In: AiML04. pp. 249–268. King's College Publications (2004)

27. Ortiz, M., Rudolph, S., Simkus, M.: Query answering in the Horn fragments of the description logics $\mathcal{SHOIQ}$ and $\mathcal{SROIQ}$. In: IJCAI. pp. 1039–1044 (2011)

28. Otto, M.: Modal and guarded characterisation theorems over finite transition systems. Ann. Pure Appl. Logic 130(1-3), 173–205 (2004)

29. Sturm, H.: Modal horn classes. Studia Logica 64(3), 301–313 (2000)

A Other Definitions of Horn Description and Modal Logics

We show that the definition of Horn-$\mathcal{ALC}$ concepts used in this paper is equivalent to syntactically different definitions of Horn-$\mathcal{ALC}$ concepts and Horn modal formulas given in the literature. To avoid confusion, in this section we refer to an $\mathcal{ALC}$ concept C which is Horn according to the definition used in this paper as a *sHorn-$\mathcal{ALC}$ concept.*

A definition for Horn DLs based on polarity is given in [17]. This definition can be restricted to the $\mathcal{ALC}$ case as follows. An $\mathcal{ALC}$ concept C is a *pHorn-$\mathcal{ALC}$ concept* if $\mathsf{pol}(C) \leq 1$, where $\mathsf{pol}(C) = \mathsf{pl}^+(C)$ with pl^+ defined as follows.

C	pl^+	pl^-
$\top$	0	0
$\bot$	0	0
A	1	0
$\neg C'$	$\mathsf{pl}^-(C')$	$\mathsf{pl}^+(C')$
$\bigsqcap_i C_i$	$\max_i \mathsf{pl}^+(C_i)$	$\sum_i \mathsf{pl}^-(C_i)$
$\bigsqcup_i C_i$	$\sum_i \mathsf{pl}^+(C_i)$	$\max_i \mathsf{pl}^-(C_i)$
$\exists r.C'$	$\max\{1, \mathsf{pl}^+(C')\}$	$\mathsf{pl}^-(C')$
$\forall r.C'$	$\mathsf{pl}^+(C')$	$\max\{1, \mathsf{pl}^-(C')\}$

Theorem 3. *Every sHorn-$\mathcal{ALC}$ concept is equivalent to a pHorn-$\mathcal{ALC}$ concept, and vice versa.*

Proof. For the left-to-right direction it is enough to show that $\mathsf{pol}(C) \leq 1$ for any sHorn-$\mathcal{ALC}$ concept C. The proof is a straightforward induction on sHorn-$\mathcal{ALC}$ concepts.

For the right-to-left direction we define a translation sH from pHorn-$\mathcal{ALC}$ concepts to equivalent sHorn-$\mathcal{ALC}$ concepts. To ease the proof, we assume the pHorn-$\mathcal{ALC}$ concepts to be in NNF. Under this assumption, the definition of pHorn-$\mathcal{ALC}$ concepts can be simplified. An $\mathcal{ALC}$ concept C in NNF is a pHorn-$\mathcal{ALC}$ concept if $\mathsf{pol}(C) \leq 1$, where $\mathsf{pol}(C)$ is defined as follows.

$$
\mathsf{pol}(C) = \begin{cases}
0 & \text{if } C = \top \mid \bot \mid \neg A \\
1 & \text{if } C = A \\
\max_i \mathsf{pol}(C_i) & \text{if } C = C_1 \sqcap \ldots \sqcap C_n \\
\sum_i \mathsf{pol}(C_i) & \text{if } C = C_1 \sqcup \ldots \sqcup C_n \\
\max\{1, \mathsf{pol}(C')\} & \text{if } C = \exists r.C' \\
\mathsf{pol}(C') & \text{if } C = \forall r.C'
\end{cases}
$$

Claim Any pHorn-$\mathcal{ALC}$ concept C in NNF of the form $C_1 \sqcup \ldots \sqcup C_n$ is equivalent to a concept of the form $L \to D$ with L an $\mathcal{ELU}$ concept, and either $D = \bot$ or $D = C_j$ for some $1 \leq j \leq n$ and $\mathsf{pol}(C_j) = 1$.

To prove the claim assume $C = C_1 \sqcup \ldots \sqcup C_n$ is a pHorn-$\mathcal{ALC}$ concept in NNF. Then $\mathsf{pol}(C) \leq 1$. It follows from the definition of pHorn that there exists at most one disjunct C_j with $1 \leq j \leq n$ and $\mathsf{pol}(C_j) = 1$. Let $C_\sqcup \sqcup D$ be a pHorn-$\mathcal{ALC}$ concept equivalent to C defined as follows. If there exists a disjunct C_j of C with $\mathsf{pol}(C_j) = 1$, then $C_\sqcup = C_1 \sqcup \ldots \sqcup C_{j-1} \sqcup C_{j+1} \sqcup \ldots \sqcup C_n$ and $D = C_j$, otherwise $C_\sqcup = C$ and $D = \bot$. It follows that $\mathsf{pol}(C_\sqcup) = 0$, which implies that $C_\sqcup$ is an $\mathcal{ALC}$ concept built using only $\top$, $\bot$, $\neg A$, $\sqcup$, $\sqcap$ and $\forall r$. Let L be the $\mathcal{ELU}$ concept defined as $\mathsf{NNF}(\neg C_\sqcup)$. Hence, C is equivalent to $L \to D$. This finishes the proof of the claim.

In the following definition of the translation sH we use L_C and D_C for a pHorn $\mathcal{ALC}$ concept $C = C_1 \sqcup \ldots \sqcup C_n$ to denote the equivalent concept $L_C \to D_C$.

$$
\mathsf{sH}(C) = \begin{cases}
C & \text{if } C ::= \top \mid \bot \mid A \mid \neg A \\
\prod_{i=1}^{n} \mathsf{sH}(C_i) & \text{if } C = C_1 \sqcap \ldots \sqcap C_n \\
L_C \to \mathsf{sH}(D_C) & \text{if } C = C_1 \sqcup \ldots \sqcup C_n \\
\exists r.\mathsf{sH}(C') & \text{if } C = \exists r.C' \\
\forall r.\mathsf{sH}(C') & \text{if } C = \forall r.C'
\end{cases}
$$

$\square$

Horn modal formulas have been defined in various ways in the literature [26, 29, 7, 8, 10], and not all definitions are equivalent. Sturm [29] defines Horn modal formulas with n-ary operators. We show that, restricted to unary modal operators, his definition is equivalent to the definition of sHorn-$\mathcal{ALC}$ concepts used in this paper. We rephrase the definition in [29] using the DL terminology as follows. Let $\mathcal{H}_b$ be defined by the rule

$$
H, H' := \bot \mid \neg A \mid H \sqcap H' \mid H \sqcup H' \mid \forall r.H
$$

where $A \in \mathsf{N_C}$ and $r \in \mathsf{N_R}$. Then the set $\mathcal{H}$ of *StHorn-$\mathcal{ALC}$ concepts* is the least set $\mathcal{H}_b \cup N_C \subseteq X$ closed under $\sqcap$, $\exists r$, and $\forall r$, and such that if $C, C' \in \mathcal{H}$ and $C \in \mathcal{H}_b$ or $C' \in \mathcal{H}_b$, then $C \sqcup C' \in \mathcal{H}$. The set $\mathcal{H}_b$ can be seen as the set containing the negation of $\mathcal{ELU}$ concepts, and thus the equivalence with our definition can be obtained by an argument analogous to the one in the proof of Theorem 3.

The remaining notions of Horn modal formulas are rather different from our definition of Horn-$\mathcal{ALC}$ concepts. To show this difference, we focus on the definition given in [26] while rephrasing it in the DL vocabulary. A *NHorn-$\mathcal{ALC}$ concept* G is defined according to the following syntax rules

$$
P, P' ::= \top \mid \bot \mid A \mid P \sqcap P' \mid P \sqcup P' \mid \exists r.P \mid \forall r.P
$$
$$
G, G' ::= A \mid \neg P \mid G \sqcap G' \mid \exists r.G \mid \forall r.G \mid P \to G
$$

The crucial difference between NHorn-$\mathcal{ALC}$ concepts and sHorn-$\mathcal{ALC}$ concepts lies in the definition of P which admits universal role restrictions $\forall r.P$. As a consequence, the concept $(\exists s.\top)\sqcap((E\sqcap\forall s.A)\to D)$ is a NHorn-$\mathcal{ALC}$ concept but not a sHorn-$\mathcal{ALC}$ concept. On the other hand, also the concept $(E\sqcap\forall s.A)\to D$ is NHorn, but it is not equivalent to any Horn FO formula.