Comparing the Expressive Powers of Rule Languages in Description Logic

Linh Anh Nguyen

Institute of Informatics, University of Warsaw,
Banacha 2, 02-097 Warsaw, Poland
nguyen@mimuw.edu.pl

Abstract. We present a method for comparing the expressive powers of rule languages in description logic. It is based on simulations w.r.t. rules, which are first introduced in this paper. Applying the method, we prove that nearly all constructors of the rule language Horn-DL are essential for its expressive power. That is, disallowing any of them decreases the expressive power of Horn-DL. We also show that the rule languages Horn-$\mathcal{SHIQ}$ and Horn-$\mathcal{SROIQ}$ and the ones with PTime data complexity of the families $\mathcal{EL}$ and DL-Lite are strictly less expressive than Horn-DL.

1 Introduction

Description logics (DLs) are useful for representing and reasoning about terminological knowledge. They found the logical base of the Web ontology language OWL. Elements of DLs are concepts, roles and individuals, where a concept is interpreted as a set of individuals and a role as a binary relation between individuals. Complex concepts and roles are built from concept names, role names and individual names using constructors. A knowledge base in DL usually consists of a finite set of axioms and assertions about roles, called an RBox, a finite set of terminological axioms, called a TBox, and a finite set of assertions about individuals, called an ABox. DLs are variants of modal logics and usually designed as decidable fragments of first-order or fixpoint logic. There is a trade-off between expressive power and complexity for DLs. OWL 2 is based on the DL $\mathcal{SROIQ}$, which is highly expressive, but also has a high combined complexity (N2ExpTime-complete) and an intractable data complexity (NP-hard).

OWL 2 RL, OWL 2 EL and OWL 2 QL are profiles of OWL 2 with PTime data complexity. OWL 2 RL was designed so that it can be translated into Datalog. OWL 2 EL and OWL 2 QL are based on the families $\mathcal{EL}$ [1,2] and DL-Lite [4] of rule languages in DL, which are Horn fragments of the corresponding full languages with appropriate restrictions adopted to eliminate nondeterminism. Other Horn fragments of DLs with PTime data complexity have also been investigated. The most notable among them are Horn-$\mathcal{SHIQ}$ [8], Horn-$\mathcal{SROIQ}$ [17] and Horn-DL [16], where Horn-DL is richer than Horn-$\mathcal{SROIQ}$, which in turn is richer than Horn-$\mathcal{SHIQ}$. In comparison with Horn-$\mathcal{SROIQ}$, Horn-DL allows a more general form for RBox axioms and additional features, including the universal role and the concept constructor $\forall\exists R.C$. This constructor stands for $\forall R.C \sqcap \exists R.C$ and is allowed to appear at the left hand side (LHS) of $\sqsubseteq$ in TBox axioms in Horn-DL, while Horn-$\mathcal{SROIQ}$ does not allow any variant of $\forall R.C$ at the LHS of $\sqsubseteq$ in TBox axioms. The authors of [16] cautiously stated only that Horn-DL is essentially richer than Horn-$\mathcal{SROIQ}$. Can we formally prove that Horn-DL is strictly more expressive than Horn-$\mathcal{SROIQ}$?

In this paper, we present a method for comparing the expressive powers of rule languages in DL. Applying the method, we prove that nearly all constructors of the rule language Horn-DL are essential for its expressive power. That is, disallowing any of them decreases the expressive power of Horn-DL. We also show that the rule languages Horn-$\mathcal{SHIQ}$ [8] and Horn-$\mathcal{SROIQ}$ [17] and the ones with PTime data complexity of the families $\mathcal{EL}$ [1,2] and DL-Lite [4] are strictly less expressive than Horn-DL.

Bisimulation has been used for analyzing the expressive power of a wide range of modal logics (see, e.g., [3]) and DLs [10,6]. It is an equivalence relation that guarantees invariance of all formulas in the sense that, if x and x' are bisimilar states, then for every formula φ, x satisfies φ iff x' satisfies φ. Simulation is a notion weaker than bisimulation (see, e.g., [3]). It is a pre-order that preserves existential formulas in the sense that, if a state x simulates x', then for every existential formula φ, if x satisfies φ, then x' also satisfies φ. Similarly, directed simulation is a notion weaker than bisimulation that preserves semi-positive formulas [9,7]. To analyze the expressive powers of rule languages, we introduce a new kind of simulation that preserves rules.

2 Preliminaries

We first recall the rule language Horn-DL [16] and its semantics. We omit the notions related with RBoxes or the clausal form as well as the instance checking problem and its data complexity. After that we define sublanguages of Horn-DL, pseudo-interpretations and comparison of rule languages w.r.t. expressive power.

2.1 The Rule Language Horn-DL

The signature consists of a finite set $\mathbf{C}$ of *concept names*, a finite set $\mathbf{R}_+$ of *role names* including a subset of *simple role names*, and a finite set $\mathbf{I}$ of *individual names*. We use letters like a, b to denote individual names, letters like A, B to denote concept names, and letters like r, s to denote role names.

For $r \in \mathbf{R}_+$, we call the expression $\overline{r}$ the *inverse* of r. Let $\mathbf{R}_- = \{\overline{r} \mid r \in \mathbf{R}_+\}$ and $\mathbf{R} = \mathbf{R}_+ \cup \mathbf{R}_-$. For $R = \overline{r}$, let $\overline{R}$ stand for r. We call elements of $\mathbf{R}$ *basic roles* and use letters like R, S to denote them. We define a *simple role* to be either a simple role name or the inverse of a simple role name.[1]

[1] Without RBoxes, any basic role can be treated as a simple role.

Concepts are defined by the following BNF grammar, where $A \in \mathbf{C}$, $R \in \mathbf{R}$, $a \in \mathbf{I}$, $n \in \mathbb{N}$ and S is a simple role:

$$\begin{aligned} C ::= &\top \mid \bot \mid A \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \forall R.C \mid \exists R.C \mid \\ &\{a\} \mid \exists R.\mathsf{Self} \mid \forall U.C \mid \exists U.C \mid \geq n\, S.C \mid \leq n\, S.C \end{aligned}$$

The symbol $U \notin \mathbf{R}$ is called the *universal role*.

We use letters like C, D to denote concepts.

Let $\forall\exists R.C$ stand for $\forall R.C \sqcap \exists R.C$. A *TBox axiom* in Horn-DL, as defined in [16], is an expression of the form $C_l \sqsubseteq C_r$, where the subscripts l and r stand for "left" and "right", respectively, C_l and C_r are concepts defined by the following BNF grammar, with $A \in \mathbf{C}$, $R \in \mathbf{R}$, $a \in \mathbf{I}$, $n \in \mathbb{N}$ and S a simple role:

$$C_l ::= \top \mid A \mid C_l \sqcap C_l \mid C_l \sqcup C_l \mid \forall\exists R.C_l \mid \exists R.C_l \mid \{a\} \mid \exists S.\mathsf{Self} \mid \forall U.C_l \mid \exists U.C_l$$

$$\begin{aligned} C_r ::= &\top \mid \bot \mid A \mid \neg C_l \mid C_r \sqcap C_r \mid \neg C_l \sqcup C_r \mid \forall R.C_r \mid \exists R.C_r \mid \forall\exists R.C_r \mid \{a\} \mid \\ &\exists R.\mathsf{Self} \mid \forall U.C_r \mid \exists U.C_r \mid \geq n\, S.C_r \mid \leq 1\, S.C_l \end{aligned}$$

Definition 2.1 (TBox and ABox in Horn-DL). A *TBox axiom* in Horn-DL can be defined equivalently as an expression of the form $C_l \sqsubseteq C_r$ using the following grammar for C_l and C_r, with $A \in \mathbf{C}$, $R \in \mathbf{R}$, $a \in \mathbf{I}$, $r \in \mathbf{R}_+$, $n \in \mathbb{N} \setminus \{0, 1\}$, s a simple role name, and S a simple role:[2]

$$\begin{aligned} C_l ::= &\top \mid A \mid C_l \sqcap C_l \mid C_l \sqcup C_l \mid \forall\exists R.C_l \mid \exists R.C_l \mid \\ &\{a\} \mid \exists s.\mathsf{Self} \mid \forall U.C_l \mid \exists U.C_l \mid \geq 2\, S.C_l \end{aligned} \tag{1}$$

$$\begin{aligned} C_r ::= &\top \mid \bot \mid A \mid \neg C_l \mid C_r \sqcap C_r \mid \neg C_l \sqcup C_r \mid \forall R.C_r \mid \exists R.C_r \mid \\ &\{a\} \mid \exists r.\mathsf{Self} \mid \forall U.C_r \mid \exists U.C_r \mid \geq n\, S.C_r \end{aligned} \tag{2}$$

A *TBox* in Horn-DL is a finite set of TBox axioms in Horn-DL.

An *ABox* in Horn-DL is a finite set of *assertions* of the form $C_r(a)$, $r(a, b)$, $\neg s(a, b)$, $a \doteq b$ or $a \not\doteq b$, where $r \in \mathbf{R}_+$, s is a simple role name and C_r is a concept defined by the grammar rule (2). □

An *interpretation* is a pair $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I} \rangle$, where $\Delta^\mathcal{I}$ is a non-empty set called the *domain* of $\mathcal{I}$ and $\cdot^\mathcal{I}$ is a mapping called the *interpretation function* of $\mathcal{I}$ that associates each individual $a \in \mathbf{I}$ with an element $a^\mathcal{I} \in \Delta^\mathcal{I}$, each concept name $A \in \mathbf{C}$ with a set $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, and each role name $r \in \mathbf{R}_+$ with a binary relation $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$. Define $\varepsilon^\mathcal{I} = \{\langle x, x \rangle \mid x \in \Delta^\mathcal{I}\}$, $U^\mathcal{I} = \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, and $(\overline{r})^\mathcal{I} = (r^\mathcal{I})^{-1} = \{\langle y, x \rangle \mid \langle x, y \rangle \in r^\mathcal{I}\}$ for $r \in \mathbf{R}_+$. The interpretation function $\cdot^\mathcal{I}$ is extended to complex concepts as shown in Figure 1. Note that, $(\forall U.C)^\mathcal{I} = \Delta^\mathcal{I}$ or $(\forall U.C)^\mathcal{I} = \emptyset$, and $(\exists U.C)^\mathcal{I} = \Delta^\mathcal{I}$ or $(\exists U.C)^\mathcal{I} = \emptyset$.

[2] One can eliminate $\neg C_l$ from (2) due to $\neg C_l \sqcup \top$. We keep it for convenience.

$$\top^{\mathcal{I}} = \Delta^{\mathcal{I}}, \quad \bot^{\mathcal{I}} = \emptyset, \quad \{a\}^{\mathcal{I}} = \{a^{\mathcal{I}}\}, \quad (\exists R.\mathsf{Self})^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \langle x, x\rangle \in R^{\mathcal{I}}\},$$
$$(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}, \quad (C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}, \quad (C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}},$$
$$(\forall R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \forall y\,(\langle x, y\rangle \in R^{\mathcal{I}} \Rightarrow y \in C^{\mathcal{I}})\},$$
$$(\exists R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \exists y\,(\langle x, y\rangle \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}})\},$$
$$(\forall U.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \Delta^{\mathcal{I}} \subseteq C^{\mathcal{I}}\}, \quad (\exists U.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid C^{\mathcal{I}} \neq \emptyset\},$$
$$(\geq n\, S.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \sharp\{y \mid \langle x, y\rangle \in S^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\} \geq n\},$$
$$(\leq n\, S.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \sharp\{y \mid \langle x, y\rangle \in S^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\} \leq n\}.$$

Fig. 1. Semantics of complex concepts.

Given an interpretation $\mathcal{I}$ and an axiom/assertion φ, the *satisfaction relation* $\mathcal{I} \models \varphi$ is defined as follows:

$$\begin{array}{ll}
\mathcal{I} \models C \sqsubseteq D & \text{if } C^{\mathcal{I}} \subseteq D^{\mathcal{I}} \\
\mathcal{I} \models C(a) & \text{if } a^{\mathcal{I}} \in C^{\mathcal{I}} \\
\mathcal{I} \models r(a, b) & \text{if } \langle a^{\mathcal{I}}, b^{\mathcal{I}}\rangle \in r^{\mathcal{I}} \\
\mathcal{I} \models \neg s(a, b) & \text{if } \langle a^{\mathcal{I}}, b^{\mathcal{I}}\rangle \notin s^{\mathcal{I}} \\
\mathcal{I} \models a \doteq b & \text{if } a^{\mathcal{I}} = b^{\mathcal{I}} \\
\mathcal{I} \models a \not\doteq b & \text{if } a^{\mathcal{I}} \neq b^{\mathcal{I}}.
\end{array}$$

If $\mathcal{I} \models \varphi$, then we say that $\mathcal{I}$ *validates* φ.

An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$ or an ABox $\mathcal{A}$ if it validates all the axioms/assertions of that "box". We write $\mathcal{I} \models \langle \mathcal{T}, \mathcal{A}\rangle$ to denote that $\mathcal{I}$ is a model of both $\mathcal{T}$ and $\mathcal{A}$. We refer the reader to [16, Example 2.3] for an illustration about Horn-DL and its semantics.

2.2 Sublanguages of Horn-DL

We name the constructors $\overline{R}$, $\forall R.C$, $\forall\exists R.C$, $\exists R.C$, $\{a\}$, $\exists r.\mathsf{Self}$, $\forall U.C$, $\exists U.C$ and $\geq n\, S.C$ by using the symbols I, $\forall$, $\forall\exists$, $\exists$, O, Self, $U_\forall$, $U_\exists$ and $Q_{\geq n}$, respectively. By a *set of features* we mean a set of these names. Let

$$\Phi_{full} = \{\forall\exists, \exists, I, O, \mathsf{Self}, U_\forall, U_\exists, Q_{\geq 2}\} \tag{3}$$
$$\Psi_{full} = \{\forall, \exists, I, O, \mathsf{Self}, U_\forall, U_\exists\} \cup \{Q_{\geq n} \mid n \geq 2\}. \tag{4}$$

A role R is a *basic role w.r.t. a set of features* if either $R = r$ or $R = \overline{r}$ and I belongs to that set, for some role name r.

Definition 2.2 (The Languages $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$). Let $\Phi \subseteq \Phi_{full}$ and $\Psi \subseteq \Psi_{full}$ be sets of features. By $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ we denote the largest sublanguage of Horn-DL such that, if a feature is absent from Φ (resp. Ψ), then the corresponding constructors are disallowed in the grammar rule (1) (resp. (2)) that defines C_l (resp. C_r). For

$$\top^\mathcal{I} = \Delta^\mathcal{I}, \quad \bot^\mathcal{I} = \emptyset, \quad \{a\}^\mathcal{I} = \{a^\mathcal{I}\}, \quad (\exists R.\mathsf{Self})^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \langle x, x\rangle \in R^{\mathcal{I}_\exists}\},$$
$$(\neg C)^\mathcal{I} = \Delta^\mathcal{I} \setminus C^\mathcal{I}, \quad (C \sqcap D)^\mathcal{I} = C^\mathcal{I} \cap D^\mathcal{I}, \quad (C \sqcup D)^\mathcal{I} = C^\mathcal{I} \cup D^\mathcal{I},$$
$$(\forall R.C)^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \forall y\,(\langle x, y\rangle \in R^{\mathcal{I}_\forall} \Rightarrow y \in C^\mathcal{I})\},$$
$$(\exists R.C)^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \exists y\,(\langle x, y\rangle \in R^{\mathcal{I}_\exists} \wedge y \in C^\mathcal{I})\},$$
$$(\forall\exists R.C)^\mathcal{I} = (\forall R.C \sqcap \exists R.C)^\mathcal{I},$$
$$(\forall U.C)^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \Delta^\mathcal{I} \subseteq C^\mathcal{I}\}, \quad (\exists U.C)^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \Delta^{\mathcal{I}_\exists} \cap C^\mathcal{I} \neq \emptyset\},$$
$$(\geq n\, S.C)^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \sharp\{y \mid \langle x, y\rangle \in S^{\mathcal{I}_\exists} \wedge y \in C^\mathcal{I}\} \geq n\},$$
$$(\leq n\, S.C)^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \sharp\{y \mid \langle x, y\rangle \in S^{\mathcal{I}_\forall} \wedge y \in C^\mathcal{I}\} \leq n\}.$$

Fig. 2. Semantics of complex concepts.

example, if $I \notin \Psi$, then when defining $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$: R and S in (2) can be only role names. As another example, if $\exists \notin \Phi$, then when defining $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ the constructor $\exists R.C_l$ is discarded from (1), while $\forall\exists R.C_l$, $\exists s.\mathsf{Self}$ and $\exists U.C_l$ are not. By a *LHS concept* (resp. *RHS concept*) of $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ we mean a concept C_l (resp. C_r) defined by the grammar for $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$. □

Note that Horn-DL is $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ with $\Phi = \Phi_{full}$ and $\Psi = \Psi_{full}$.

2.3 Pseudo-Interpretations

The notion of pseudo-interpretation was introduced by us in [12,13,14,15]. The following definition uses the notations of [15], but is closer to [14].

Definition 2.3 (Pseudo-Interpretation). A *pseudo-interpretation* is a pair $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I}\rangle$, where $\Delta^\mathcal{I}$ is a non-empty set called the *domain* of $\mathcal{I}$ and $\cdot^\mathcal{I}$ is a mapping called the *interpretation function* of $\mathcal{I}$ that associates each individual name $a \in \mathbf{I}$ with an element $a^\mathcal{I} \in \Delta^\mathcal{I}$, each concept name $A \in \mathbf{C}$ with a set $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, and each role name $r \in \mathbf{R}_+$ with a pair $\langle r^{\mathcal{I}_\exists}, r^{\mathcal{I}_\forall}\rangle$ of binary relations such that $r^{\mathcal{I}_\exists} \subseteq r^{\mathcal{I}_\forall} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$. The interpretation function also specifies a set $\Delta^{\mathcal{I}_\exists} \subseteq \Delta^\mathcal{I}$. We define $\varepsilon^{\mathcal{I}_\exists} = \varepsilon^{\mathcal{I}_\forall} = \{\langle x, x\rangle \mid x \in \Delta^\mathcal{I}\}$, $U^{\mathcal{I}_\exists} = \Delta^\mathcal{I} \times \Delta^{\mathcal{I}_\exists}$, $U^{\mathcal{I}_\forall} = \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, $(\overline{r})^{\mathcal{I}_\exists} = (r^{\mathcal{I}_\exists})^{-1}$ and $(\overline{r})^{\mathcal{I}_\forall} = (r^{\mathcal{I}_\forall})^{-1}$ for $r \in \mathbf{R}_+$. The interpretation function $\cdot^\mathcal{I}$ is extended to complex concepts as shown in Figure 2. □

Remark 2.4. An interpretation $\mathcal{I}$ can be treated as a pseudo-interpretation with $\Delta^{\mathcal{I}_\exists} = \Delta^\mathcal{I}$ and $r^{\mathcal{I}_\exists} = r^{\mathcal{I}_\forall} = r^\mathcal{I}$ for all $r \in \mathbf{R}_+$. Conversely, a pseudo-interpretation $\mathcal{I}$ satisfying $\Delta^{\mathcal{I}_\exists} = \Delta^\mathcal{I}$ and $r^{\mathcal{I}_\exists} = r^{\mathcal{I}_\forall} = r^\mathcal{I}$ for all $r \in \mathbf{R}_+$ can be treated as an interpretation. □

Observe that, given a pseudo-interpretation $\mathcal{I}$ and a role R, we have that $R^{\mathcal{I}_\exists} \subseteq R^{\mathcal{I}_\forall}$, and $(\forall R.C)^\mathcal{I}$ may differ from $(\neg\exists R.\neg C)^\mathcal{I}$. In addition, $(\forall U.C)^\mathcal{I}$

may differ from $(\neg\exists U.\neg C)^{\mathcal{I}}$. If $\langle x,y\rangle \in R^{\mathcal{I}_\exists}$, then we call $\langle x,y\rangle$ a *firm R-edge*. If $\langle x,y\rangle \in R^{\mathcal{I}_\forall} \setminus R^{\mathcal{I}_\exists}$, then we call $\langle x,y\rangle$ a *pseudo R-edge*. If $x \in \Delta^{\mathcal{I}}$, then we call x a *firm individual*. If $x \in \Delta^{\mathcal{I}} \setminus \Delta^{\mathcal{I}_\exists}$, then we call x a *pseudo individual*. For duality, we can use $\Delta^{\mathcal{I}_\forall} = \Delta^{\mathcal{I}}$ and treat $U^{\mathcal{I}_\forall}$ as $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}_\forall}$.

Given a pseudo-interpretation $\mathcal{I}$ and an axiom/assertion φ, the *satisfaction relation* $\mathcal{I} \models\!\!\!\!\sim\, \varphi$ is defined as follows:

$$\begin{array}{ll}
\mathcal{I} \models\!\!\!\!\sim C \sqsubseteq D & \text{if } C^{\mathcal{I}} \subseteq D^{\mathcal{I}} \\
\mathcal{I} \models\!\!\!\!\sim C(a) & \text{if } a^{\mathcal{I}} \in C^{\mathcal{I}} \\
\mathcal{I} \models\!\!\!\!\sim r(a,b) & \text{if } \langle a^{\mathcal{I}}, b^{\mathcal{I}}\rangle \in r^{\mathcal{I}_\exists} \\
\mathcal{I} \models\!\!\!\!\sim \neg s(a,b) & \text{if } \langle a^{\mathcal{I}}, b^{\mathcal{I}}\rangle \notin s^{\mathcal{I}_\forall} \\
\mathcal{I} \models\!\!\!\!\sim a \doteq b & \text{if } a^{\mathcal{I}} = b^{\mathcal{I}} \\
\mathcal{I} \models\!\!\!\!\sim a \not\doteq b & \text{if } a^{\mathcal{I}} \neq b^{\mathcal{I}}.
\end{array}$$

If $\mathcal{I} \models\!\!\!\!\sim \varphi$, then we say that $\mathcal{I}$ *validates* φ. A pseudo-interpretation $\mathcal{I}$ is a *pseudo-model* of a TBox $\mathcal{T}$ or an ABox $\mathcal{A}$ if it validates all the axioms/assertions of that "box". We write $\mathcal{I} \models\!\!\!\!\sim \langle \mathcal{T}, \mathcal{A}\rangle$ to denote that $\mathcal{I}$ is a pseudo-model of both $\mathcal{T}$ and $\mathcal{A}$.

2.4 On the Expressive Power of Rule Languages

In this subsection, we define a notion for comparing the expressive powers of rule languages and then justify that definition.

Definition 2.5 (Comparison w.r.t. Expressive Power). Let L be a sublanguage of a language L' in DL (e.g., L' = Horn-DL and L is some $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$). We say that L is *strictly less expressive* than L', denoted by $L < L'$, if there exists a TBox $\mathcal{T}'$ and an ABox $\mathcal{A}'$ in L' such that, for every TBox $\mathcal{T}$ and every ABox $\mathcal{A}$ in L, there exists a pseudo-interpretation $\mathcal{I}$ such that $\mathcal{I} \models\!\!\!\!\sim \langle \mathcal{T}, \mathcal{A}\rangle$ iff $\mathcal{I} \not\models\!\!\!\!\sim \langle \mathcal{T}', \mathcal{A}'\rangle$. If φ is a TBox axiom (resp. ABox assertion) in L' and there do not exist any TBox $\mathcal{T}$ and any ABox $\mathcal{A}$ in L such that, for every pseudo-interpretation $\mathcal{I}$, $\mathcal{I} \models\!\!\!\!\sim \langle \mathcal{T}, \mathcal{A}\rangle$ iff $\mathcal{I} \models\!\!\!\!\sim \langle \{\varphi\}, \emptyset\rangle$ (resp. $\mathcal{I} \models\!\!\!\!\sim \langle \emptyset, \{\varphi\}\rangle$), then we say that φ is *inexpressible* in L. □

To justify this definition, we first explain the intuition behind pseudo-interpretations. Suppose that $C \sqsubseteq \forall R.D$ is an axiom of $\mathcal{T}$ and when constructing a least or minimal model for $\langle \mathcal{T}, \mathcal{A}\rangle$ we have an individual x satisfying C. Naturally, we add $\forall R.D$ to the label of x as a requirement for it. How can we realize this requirement? One would add D to the label of every R-successor of x. Is it enough? Suppose $\forall R.D' \sqsubseteq E$ is another axiom of $\mathcal{T}$. To satisfy this axiom for x, the question is "how to check whether x satisfies $\forall R.D'$ or not?". Checking whether D' currently holds for every R-successor of x is not a sufficient solution for that problem. The reason is that D' may accidentally hold for every R-successor of x, but $\forall R.D'$ does not really follow from the requirements for x. A solution for the raised problems is as follows. For each individual x, we connect

x to its *least R-successor*, let's call it y, which is created when necessary. Realizing a requirement $\forall R.D$ for x, D is also added to the label of this successor y of x. Then, to check whether $\forall R.D'$ holds for x, it is sufficient to check whether D' holds for y. However, this solution causes an unexpected side effect that $\exists R.\top$ holds for x, despite the fact that it may not follow from the requirements for x. So, y should only be a *pseudo R-successor of* x. That is, $\langle x, y\rangle$ should only be a *pseudo R-edge*. It is taken into account when dealing with concepts of the form $\forall R.F$, but ignored when dealing with concepts of the form $\exists R.F$.

The intuition behind pseudo individuals is similar. When constructing a least or minimal model for $\langle \mathcal{T}, \mathcal{A}\rangle$, checking whether $\forall U.C$ holds for x cannot simply base on checking whether C holds for all individuals. Our solution is to create and use the *least individual*. It is taken into account when dealing with concepts of the form $\forall U.C$, but is not essential when dealing with concepts of the form $\exists U.C$. The least individual is a *pseudo individual*.

As explained above, pseudo-interpretations are a natural notion for dealing with rule languages. They have been used, among others, in [12,13,14,15]. So, it makes sense to use pseudo-interpretations for studying the expressive powers of rule languages. We discuss below three other aspects of Definition 2.5:

- Our definition of whether an axiom/assertion φ is inexpressible in L is liberal in the sense that we allow to use a TBox and an ABox in L to express it.
- We ignore RBoxes because we want to concentrate on analyzing the influence of concept and role constructors on the expressive powers of rule languages. Axioms and assertions about roles are usually formulated for basic roles, without using constructors except the one for inverse roles. If one wants to incorporate RBoxes into Definition 2.5, for a knowledge base $\langle \mathcal{R}', \mathcal{T}', \mathcal{A}'\rangle$ in L' with $\mathcal{R}'$ being an RBox, he or she can look for an equivalent knowledge base $\langle \mathcal{R}, \mathcal{T}, \mathcal{A}\rangle$ such that $\mathcal{T}$ and $\mathcal{A}$ are in L and $\mathcal{R} = \mathcal{R}'$. Our definition is liberal in this aspect because $\mathcal{R}$ is allowed to be in L', a superlanguage of L.
- We use the same signature ($\mathbf{C}$, $\mathbf{R}_+$ and $\mathbf{I}$) for both L and L'. This is a natural approach for comparing languages w.r.t. the influence of constructors on the expressive power. A more liberal approach is to allow L to use a richer signature. Additional concept names are usually used as macros that refer to concepts defined in L using the original signature. We overcome this problem by considering rich rule languages that can eliminate such macros. In particular, Horn-DL and its sublanguages are considered in the general form, but not restricted to the clausal form [16].

3 Simulation with Respect to Rules

In the following, given a set P, we will write $P(x)$ to denote that $x \in P$. Similarly, given a binary relation P, we will write $P(x, y)$ to denote that $\langle x, y\rangle \in P$.

Definition 3.1 (Φ-Simulation between Pseudo-Interpretations). Let $\Phi \subseteq \Phi_{full}$ and let $\mathcal{I}$ and $\mathcal{I}'$ be pseudo-interpretations. A non-empty binary relation $Z \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}'}$ is called a *Φ-simulation* between $\mathcal{I}$ and $\mathcal{I}'$ if the following conditions hold for every $a \in \Sigma_I$, $A \in \mathbf{C}$, every basic role R w.r.t. Φ, every simple role name s and every $x, y \in \Delta^{\mathcal{I}}$, $x', y' \in \Delta^{\mathcal{I}'}$:

$$Z(a^{\mathcal{I}}, a^{\mathcal{I}'}) \tag{5}$$

$$Z(x, x') \Rightarrow [A^{\mathcal{I}}(x) \Rightarrow A^{\mathcal{I}'}(x')]; \tag{6}$$

if $\exists \in \Phi$, then

$$[Z(x, x') \wedge R^{\mathcal{I}_\exists}(x, y)] \Rightarrow \exists y' \, [Z(y, y') \wedge R^{\mathcal{I}'_\exists}(x', y')]; \tag{7}$$

if $\forall\exists \in \Phi$, then

$$[Z(x, x') \wedge R^{\mathcal{I}_\exists}(x, y)] \Rightarrow \exists y' \, R^{\mathcal{I}'_\exists}(x', y') \tag{8}$$

$$[Z(x, x') \wedge R^{\mathcal{I}'_\forall}(x', y') \wedge \exists y \, R^{\mathcal{I}_\exists}(x, y)] \Rightarrow \exists y \, [Z(y, y') \wedge R^{\mathcal{I}_\forall}(x, y)]; \tag{9}$$

if $O \in \Phi$, then

$$Z(x, x') \Rightarrow [x = a^{\mathcal{I}} \Rightarrow x' = a^{\mathcal{I}'}]; \tag{10}$$

if $Q_{\geq 2} \in \Phi$, then

if $Z(x, x')$ holds and y_1, y_2 are different elements of $\Delta^{\mathcal{I}}$ such that both $R^{\mathcal{I}_\exists}(x, y_1)$ and $R^{\mathcal{I}_\exists}(x, y_2)$ hold, then there exist different elements y'_1 and y'_2 of $\Delta^{\mathcal{I}'}$ such that, for each $i \in \{1, 2\}$, $R^{\mathcal{I}'_\exists}(x', y'_i)$ and $Z(y_j, y'_i)$ hold for some $j \in \{1, 2\}$; (11)

if $U_\exists \in \Phi$, then

$$\forall x \in \Delta^{\mathcal{I}_\exists} \, \exists x' \in \Delta^{\mathcal{I}'_\exists} \, Z(x, x'); \tag{12}$$

if $U_\forall \in \Phi$, then

$$\forall x' \in \Delta^{\mathcal{I}'} \, \exists x \in \Delta^{\mathcal{I}} \, Z(x, x'); \tag{13}$$

if $\mathsf{Self} \in \Phi$, then

$$Z(x, x') \Rightarrow [s^{\mathcal{I}_\exists}(x, x) \Rightarrow s^{\mathcal{I}'_\exists}(x', x')]. \tag{14}$$

We write $\mathcal{I} \lesssim_\Phi \mathcal{I}'$ to denote that there exists a Φ-simulation between $\mathcal{I}$ and $\mathcal{I}'$. For $x \in \Delta^{\mathcal{I}}$ and $x' \in \Delta^{\mathcal{I}'}$, we write $x \lesssim_\Phi x'$ to denote that there exists a Φ-simulation Z between $\mathcal{I}$ and $\mathcal{I}'$ such that $Z(x, x')$ holds. □

Definition 3.2 ((Φ, Ψ)-Simulation between Pseudo-Interpretations). Let $\Phi \subseteq \Phi_{full}$, $\Psi \subseteq \Psi_{full}$ and let $\mathcal{I}$ and $\mathcal{I}'$ be interpretations. A non-empty binary relation $Z \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}'}$ is called a (Φ, Ψ)-*simulation* between $\mathcal{I}$ and $\mathcal{I}'$ if the following conditions hold for every $a \in \Sigma_I$, $A \in \mathbf{C}$, $r \in \mathbf{R}_+$, every basic role R w.r.t. Ψ and every $x, y \in \Delta^{\mathcal{I}}$, $x', y' \in \Delta^{\mathcal{I}'}$:

$$Z(a^{\mathcal{I}}, a^{\mathcal{I}'}) \tag{15}$$

$$Z(x, x') \Rightarrow [A^{\mathcal{I}}(x) \Rightarrow A^{\mathcal{I}'}(x')] \tag{16}$$

$$Z(x, x') \Rightarrow x' \lesssim_\Phi x; \tag{17}$$

if $\exists \in \Psi$, then

$$[Z(x,x') \land R^{\mathcal{I}_\exists}(x,y)] \Rightarrow \exists y' \, [Z(y,y') \land R^{\mathcal{I}'_\exists}(x',y')]; \tag{18}$$

if $\forall \in \Psi$, then

$$[Z(x,x') \land R^{\mathcal{I}'_\forall}(x',y')] \Rightarrow \exists y \, [Z(y,y') \land R^{\mathcal{I}_\forall}(x,y)]; \tag{19}$$

if $O \in \Psi$, then

$$Z(x,x') \Rightarrow [x = a^{\mathcal{I}} \Rightarrow x' = a^{\mathcal{I}'}]; \tag{20}$$

if $Q_{\geq n} \in \Psi$, then

> if $Z(x,x')$ holds and $y_1, \ldots, y_n$ are pairwise distinct elements of $\Delta^{\mathcal{I}}$ such that $R^{\mathcal{I}_\exists}(x,y_i)$ holds for every $1 \leq i \leq n$, then there exist pairwise distinct elements $y'_1, \ldots, y'_n$ of $\Delta^{\mathcal{I}'}$ such that, for every $1 \leq i \leq n$, $R^{\mathcal{I}'_\exists}(x',y'_i)$ and $Z(y_j,y'_i)$ hold for some $1 \leq j \leq n$; (21)

if $U_\exists \in \Psi$, then

$$\forall x \in \Delta^{\mathcal{I}_\exists} \, \exists x' \in \Delta^{\mathcal{I}'_\exists} \, Z(x,x'); \tag{22}$$

if $U_\forall \in \Psi$, then

$$\forall x' \in \Delta^{\mathcal{I}'} \, \exists x \in \Delta^{\mathcal{I}} \, Z(x,x'); \tag{23}$$

if $\mathsf{Self} \in \Psi$, then

$$Z(x,x') \Rightarrow [r^{\mathcal{I}_\exists}(x,x) \Rightarrow r^{\mathcal{I}'_\exists}(x',x')]. \tag{24}$$

We write $\mathcal{I} \lesssim_{\Phi,\Psi} \mathcal{I}'$ to denote that there exists a (Φ,Ψ)-simulation between $\mathcal{I}$ and $\mathcal{I}'$. For $x \in \Delta^{\mathcal{I}}$ and $x' \in \Delta^{\mathcal{I}'}$, we write $x \lesssim_{\Phi,\Psi} x'$ to denote that there exists a (Φ,Ψ)-simulation Z between $\mathcal{I}$ and $\mathcal{I}'$ such that $Z(x,x')$ holds. □

We say that a concept C is *preserved by Φ-simulations* (resp. *(Φ,Ψ)-simulations*) if, for any pseudo-interpretations $\mathcal{I}$, $\mathcal{I}'$ and any Φ-simulation (resp. (Φ,Ψ)-simulation) Z between $\mathcal{I}$ and $\mathcal{I}'$, if $Z(x,x')$ holds and $x \in C^{\mathcal{I}}$, then $x' \in C^{\mathcal{I}'}$.

Theorem 3.3. *All LHS concepts of $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ are preserved by Φ-simulations. All RHS concepts of $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ are preserved by (Φ,Ψ)-simulations.*

This theorem can be proved in a similar way as for [7, Theorem 5.3]. For the case when C is a LHS concept of the form $\forall\exists R.D$, note that $(\forall\exists R.D)^{\mathcal{I}} = (\forall R.D \sqcap \exists R.\top)^{\mathcal{I}}$ for any pseudo-interpretation $\mathcal{I}$. When C is a RHS concept of the form $\neg C_l$ we have an additional base step of the induction. It uses Condition (17).

We say that a TBox $\mathcal{T}$ is *preserved by (Φ,Ψ)-simulations* if, for every pseudo-interpretations $\mathcal{I}$ and $\mathcal{I}'$ and every (Φ,Ψ)-simulation Z between $\mathcal{I}$ and $\mathcal{I}'$, if $Z(x,x')$ holds and $\mathcal{I} \models \mathcal{T}$, then $\mathcal{I}' \models \mathcal{T}$. The notion of whether an ABox is preserved by (Φ,Ψ)-simulations is defined similarly.

We present below propositions about conditional preservation of TBoxes and ABoxes by (Φ,Ψ)-simulations. We refer the reader to the full version [11] of this paper for their proofs.

Proposition 3.4. *If $U_\forall \in \Psi$, then all TBoxes in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ are preserved by (Φ, Ψ)-simulations.*

Proposition 3.5. *Let $\mathcal{T}$ be a TBox in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ and $\mathcal{I}, \mathcal{I}'$ be pseudo-interpretations. If $\mathcal{I} \models \mathcal{T}$ and, for every $x' \in \Delta^{\mathcal{I}'}$, there exists $x \in \Delta^{\mathcal{I}}$ such that $x \lesssim_{\Phi,\Psi} x'$, then $\mathcal{I}' \models \mathcal{T}$.*

A pseudo-interpretation $\mathcal{I}$ is said to be *connected* (or *unreachable-objects-free* [6]) w.r.t. a set Ψ of features if, for every $x \in \Delta^{\mathcal{I}}$, there exist $x_0, \ldots, x_k \in \Delta^{\mathcal{I}}$ and basic roles $R_1, \ldots, R_k$ w.r.t. Ψ such that $k \geq 0$, $x_0 = a^{\mathcal{I}}$ for some $a \in \mathbf{I}$, $x_k = x$ and $R_i^{\mathcal{I}_\forall}(x_{i-1}, x_i)$ holds for every $1 \leq i \leq k$.

Proposition 3.6. *Let $\mathcal{T}$ be a TBox in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ and $\mathcal{I}, \mathcal{I}'$ be pseudo-interpretations. If $\mathcal{I} \lesssim_{\Phi,\Psi} \mathcal{I}'$, $\mathcal{I} \models \mathcal{T}$, $\forall \in \Psi$ and $\mathcal{I}'$ is connected w.r.t. Ψ, then $\mathcal{I}' \models \mathcal{T}$.*

Proposition 3.7. *Let $\mathcal{I}$ and $\mathcal{I}'$ be pseudo-interpretations such that $\mathcal{I} \lesssim_{\Phi,\Psi} \mathcal{I}'$ and, for every assertion φ of the form $r(a, b)$, $\neg s(a, b)$, $a \doteq b$ or $a \not\doteq b$, if $\mathcal{I} \models \varphi$, then $\mathcal{I}' \models \varphi$. Then, for every ABox $\mathcal{A}$ in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$, if $\mathcal{I} \models \mathcal{A}$, then $\mathcal{I}' \models \mathcal{A}$.*

Proposition 3.8. *Let $\mathcal{A}$ be an ABox in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ such that:*

- *if $\mathcal{A}$ contains an assertion of the form $r(a, b)$, then $O \in \Psi$ and $\exists \in \Psi$;*
- *if $\mathcal{A}$ contains an assertion of the form $\neg s(a, b)$, then $O \in \Phi$ and $\forall \in \Psi$;*
- *if $\mathcal{A}$ contains an assertion of the form $a \doteq b$, then $O \in \Psi$;*
- *if $\mathcal{A}$ contains an assertion of the form $a \not\doteq b$, then $O \in \Phi$.*

Then $\mathcal{A}$ is preserved by (Φ, Ψ)-simulations.

4 On the Expressive Power of Horn-DL

We will consider the languages $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ such that either $\Phi = \Phi_{full}$ and Ψ is obtained from Ψ_{full} by deleting one feature, or conversely, $\Psi = \Psi_{full}$ and Φ is obtained from Φ_{full} by deleting one feature. For one case, we show that $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ has the same expressive power as Horn-DL. For all the other cases, we show that $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ is strictly less expressive than Horn-DL. The proof for these cases has the following pattern. We choose a signature $(\mathbf{C}, \mathbf{R}_+, \mathbf{I})$ and two pseudo-interpretations $\mathcal{I}$ and $\mathcal{I}'$ together with an axiom/assertion φ of Horn-DL. We then show that $\mathcal{I} \lesssim_{\Phi,\Psi} \mathcal{I}'$, $\mathcal{I} \models \varphi$, but $\mathcal{I}' \not\models \varphi$. This implies that φ is inexpressible in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$.

Due to the lack of space, we present details only for the cases when $\Phi = \Phi_{full} \setminus \{\forall\exists\}$ or $\Psi = \Psi_{full} \setminus \{\forall\}$ and refer the reader to the full version [11] of this paper for the other cases. The following theorem follows from Lemmas 4.3 and 4.4 and [11, Lemmas 4.5–4.18].

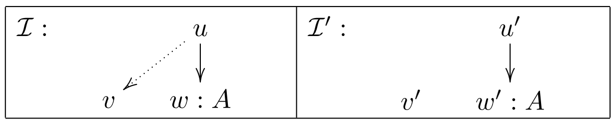
Theorem 4.1. *Let $\Phi \subseteq \Phi_{full}$ and $\Psi \subseteq \Psi_{full}$. If $\Phi \subset \Phi_{full}$ or $\Psi \cup \{U_\forall\} \subset \Psi_{full}$, then the language $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ is strictly less expressive than Horn-DL. If $\Phi = \Phi_{full}$ and $\Psi = \Psi_{full} \setminus \{U_\forall\}$, then $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ has the same expressive power as Horn-DL.*

Corollary 4.2. *The rule languages Horn-$\mathcal{SHIQ}$ [8] and Horn-$\mathcal{SROIQ}$ [17] and the ones with* PTime *data complexity of the families $\mathcal{EL}$ [1,2] and DL-Lite [4] are strictly less expressive than Horn-DL.*

This corollary holds because all the mentioned rule languages (except Horn-DL) are sublanguages of $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ with $\Phi = \Phi_{full} \setminus \{\forall\exists\}$ and $\Psi = \Psi_{full}$.

4.1 The Case of ∀∃ on the LHS

Let $\mathbf{C} = \{A\}$, $\mathbf{R}_+ = \{r\}$, $\mathbf{I} = \emptyset$. Consider the following pseudo-interpretations:



- $\Delta^{\mathcal{I}} = \Delta^{\mathcal{I}_\exists} = \{u, v, w\}$, $A^{\mathcal{I}} = \{w\}$, $r^{\mathcal{I}_\exists} = \{\langle u, w\rangle\}$, $r^{\mathcal{I}_\forall} = r^{\mathcal{I}_\exists} \cup \{\langle u, v\rangle\}$,
- $\Delta^{\mathcal{I}'} = \Delta^{\mathcal{I}'_\exists} = \{u', v', w'\}$, $A^{\mathcal{I}'} = \{w'\}$, $r^{\mathcal{I}'_\forall} = r^{\mathcal{I}'_\exists} = \{\langle u', v'\rangle, \langle u', w'\rangle\}$.

Let $\Phi = \Phi_{full} \setminus \{\forall\exists\}$ and $\Psi = \Psi_{full}$. Observe that:

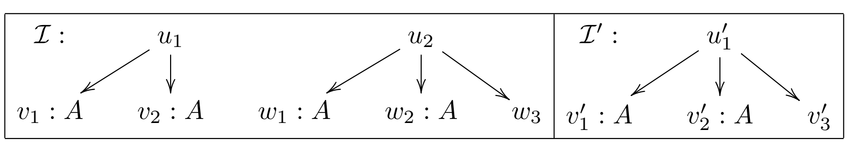
- $Z' = \{\langle u', u\rangle, \langle v', v\rangle, \langle w', w\rangle\}$ is a Φ-simulation between $\mathcal{I}'$ and $\mathcal{I}$,
- $Z = (Z')^{-1}$ is a (Φ, Ψ)-simulation between $\mathcal{I}$ and $\mathcal{I}'$,
- $\mathcal{I} \models (\forall\exists r.A \sqsubseteq A)$, but $\mathcal{I}' \not\models (\forall\exists r.A \sqsubseteq A)$.

Lemma 4.3. *The axiom $\forall\exists r.A \sqsubseteq A$ of Horn-DL is inexpressible in the language $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ with $\Phi = \Phi_{full} \setminus \{\forall\exists\}$ and $\Psi = \Psi_{full}$.*

Proof. Let $\Phi = \Phi_{full} \setminus \{\forall\exists\}$ and $\Psi = \Psi_{full}$. Consider $\mathcal{T} = \{\forall\exists r.A \sqsubseteq A\}$, $\mathcal{A} = \emptyset$ and the pseudo-interpretations $\mathcal{I}, \mathcal{I}'$ specified in this subsection. For a contradiction, suppose $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ has the same expressive power as Horn-DL. Thus, there exist a TBox $\mathcal{T}'$ and an ABox $\mathcal{A}'$ in $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ such that, for every pseudo-interpretation $\mathcal{I}''$, $\mathcal{I}'' \models \langle \mathcal{T}, \mathcal{A}\rangle$ iff $\mathcal{I}'' \models \langle \mathcal{T}', \mathcal{A}'\rangle$. Since $\mathcal{I} \models \langle \mathcal{T}, \mathcal{A}\rangle$, we have $\mathcal{I} \models \langle \mathcal{T}', \mathcal{A}'\rangle$. Since $\mathcal{I} \lesssim_{\Phi,\Psi} \mathcal{I}'$, by Propositions 3.4 and 3.8, it follows that $\mathcal{I}' \models \langle \mathcal{T}', \mathcal{A}'\rangle$, which implies $\mathcal{I}' \models \langle \mathcal{T}, \mathcal{A}\rangle$, a contradiction. □

4.2 The Case of ∀ on the RHS

Let $\mathbf{C} = \{A\}$, $\mathbf{R}_+ = \{r\}$ and $\mathbf{I} = \emptyset$. Consider the following interpretations:



- $\Delta^{\mathcal{I}} = \{u_1, u_2, v_1, v_2, w_1, w_2, w_3\}$, $A^{\mathcal{I}} = \{v_1, v_2, w_1, w_2\}$,
- $r^{\mathcal{I}} = \{\langle u_1, v_i\rangle, \langle u_2, w_j\rangle \mid 1 \le i \le 2,\ 1 \le j \le 3\}$,
- $\Delta^{\mathcal{I}'} = \{u'_1, v'_1, v'_2, v'_3\}$, $A^{\mathcal{I}'} = \{v'_1, v'_2\}$, $r^{\mathcal{I}'} = \{\langle u'_1, v'_i\rangle \mid 1 \le i \le 3\}$.

We treat $\mathcal{I}$ and $\mathcal{I}'$ as pseudo-interpretations. Let $\Phi = \Phi_{full}$ and $\Psi = \Psi_{full} \setminus \{\forall\}$. Observe that:

- $Z' = \{\langle u'_1, u_1\rangle, \langle u'_1, u_2\rangle, \langle v'_h, v_i\rangle, \langle v'_h, w_j\rangle \mid \{h, i, j\} \subseteq \{1, 2, 3\}, i \leq 2, (j = 3 \Rightarrow h = 3)\}$ is a Φ-simulation between $\mathcal{I}'$ and $\mathcal{I}$,
- $Z = \{\langle u_1, u'_1\rangle, \langle u_2, u'_1\rangle, \langle v_i, v'_i\rangle, \langle w_j, v'_j\rangle \mid 1 \leq i \leq 2, 1 \leq j \leq 3\}$ is a (Φ, Ψ)-simulation between $\mathcal{I}$ and $\mathcal{I}'$,
- $\mathcal{I} \models (\top \sqsubseteq \exists U.(\forall r.A \sqcap \exists r.\top)$, but $\mathcal{I}' \not\models (\top \sqsubseteq \exists U.(\forall r.A \sqcap \exists r.\top))$.

These observations imply the following lemma, whose proof is similar to the one of Lemma 4.3 (but uses Proposition 3.5 instead of Proposition 3.4).

Lemma 4.4. *The axiom $\top \sqsubseteq \exists U.(\forall r.A \sqcap \exists r.\top)$ of Horn-DL is inexpressible in the language $\mathcal{L}^{\sqsubseteq}_{\Phi,\Psi}$ with $\Phi = \Phi_{full}$ and $\Psi = \Psi_{full} \setminus \{\forall\}$.*

5 Discussion and Conclusions

We have provided a method for comparing the expressive powers of rule languages in description logic. It is based on simulations w.r.t. rules, which are first introduced in this paper. Such simulations are a generalization of directed simulations [9,7] and allow us to deal with rules instead of semi-positive concepts. A fundamental result related with such simulations is the preservation of rules (Theorem 3.3). Some other results concern conditional preservation of TBoxes and ABoxes by such simulations (Propositions 3.4–3.8). Our formulation of simulations w.r.t. rules is loose in the sense that it can be used for monotonic disjunctive rules (i.e., one can allow the pattern $C_r \sqcup C_r$ instead of $\neg C_l \sqcup C_r$ in the grammar rule defining C_r like (2)), and using such simulations one can prove the Hennessy-Milner property only for monotonic disjunctive rules. We chose that formulation to increase simplicity, knowing that our simulations w.r.t. rules are sufficient to separate a large class of rule languages w.r.t. expressive power.

Applying the method to Horn-DL and its sublanguages, we have obtained Theorem 4.1 and Corollary 4.2, which are nice theoretical results.

As discussed in Section 2.4, our definition of comparing the expressive powers of rule languages is acceptable and liberal. In particular, using the same signature for the compared rule languages is natural. In [5], Carral et al. proposed a method for transforming a knowledge base KB in the rule language that extends Horn-$\mathcal{SROIQ}$ with the constructor $\forall\exists R.C$ at the LHS of $\sqsubseteq$ in TBox axioms to another knowledge base KB' in Horn-$\mathcal{SROIQ}$ such that KB and KB' are equisatisfiable and KB' does not use the constructor $\forall\exists R.C$ at the LHS of $\sqsubseteq$ in TBox axioms. It is based on a technique used in our algorithms for checking satisfiability of a Horn knowledge base in DLs [12,16]. Technically, the transformation uses additional concept names, role names and RBox axioms. The question is: does it mean that the constructor $\forall\exists R.C$ at the LHS of $\sqsubseteq$ in TBox axioms does not increase the expressive powers of Horn-$\mathcal{SROIQ}$ and Horn-DL? The answer is "no". First, we have Theorem 4.1 for the setting without using additional role names. Second, the mentioned equisatisfiability is too weak to deal with expressive power. Namely, one can check satisfiability of KB and, depending on the result, take either an empty knowledge base or a simple unsatisfiable knowledge base for KB'.

References

1. F. Baader, S. Brandt, and C. Lutz. Pushing the EL envelope. In L.P. Kaelbling and A. Saffiotti, editors, *Proceedings of IJCAI'2005*, pages 364–369. Morgan-Kaufmann Publishers, 2005.
2. F. Baader, S. Brandt, and C. Lutz. Pushing the EL envelope further. In *Proceedings of the OWLED 2008 DC Workshop on OWL: Experiences and Directions*, 2008.
3. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Number 53 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2001.
4. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The DL-Lite family. *J. Autom. Reasoning*, 39(3):385–429, 2007.
5. D. Carral, A. Krisnadhi, S. Rudolph, and P. Hitzler. All but not nothing: Left-hand side universals for tractable OWL profiles. In *Proceedings of OWLED 2014*, volume 1265 of *CEUR Workshop Proceedings*, pages 97–108. CEUR-WS.org, 2014.
6. A.R. Divroodi and L.A. Nguyen. On bisimulations for description logics. *Information Sciences*, 295:465–493, 2015.
7. A.R. Divroodi and L.A. Nguyen. On directed simulations in description logics. *Journal of Logic and Computation*, doi:`10.1093/logcom/exx004`, 2017.
8. U. Hustadt, B. Motik, and U. Sattler. Reasoning in description logics by a reduction to disjunctive Datalog. *J. Autom. Reasoning*, 39(3):351–384, 2007.
9. N. Kurtonina and M. de Rijke. Simulating without negation. *J. Log. Comput.*, 7(4):501–522, 1997.
10. N. Kurtonina and M. de Rijke. Expressiveness of concept expressions in first-order description logics. *Artif. Intell.*, 107(2):303–333, 1999.
11. L.A. Nguyen. A long version of the current paper. `http://www.mimuw.edu.pl/~nguyen/EP-RL-DL-long.pdf`.
12. L.A. Nguyen. A bottom-up method for the deterministic Horn fragment of the description logic $\mathcal{ALC}$. In *Proceedings of JELIA'2006*, volume 4160 of *LNAI*, pages 346–358. Springer-Verlag, 2006.
13. L.A. Nguyen. On the deterministic Horn fragment of test-free PDL. In I. Hodkinson and Y. Venema, editors, *Advances in Modal Logic - Volume 6*, pages 373–392. King's College Publications, 2006.
14. L.A. Nguyen. Horn knowledge bases in regular description logics with PTime data complexity. *Fundamenta Informaticae*, 104(4):349–384, 2010.
15. L.A. Nguyen. A tractable rule language in the modal and description logic that combines CPDL with regular grammar logic. *Fundam. Inform.*, 143(1-2):113–139, 2016.
16. L.A. Nguyen, T.-B.-L. Nguyen, and A. Szałas. Towards richer rule languages with polynomial data complexity for the Semantic Web. *Data Knowl. Eng.*, 96:57–77, 2015.
17. M. Ortiz, S. Rudolph, and M. Simkus. Query answering in the Horn fragments of the description logics SHOIQ and SROIQ. In T. Walsh, editor, *Proceedings of IJCAI 2011*, pages 1039–1044, 2011.